IDA

INSTITUTE FOR DEFENSE ANALYSES

Comparisons of Transport and Dispersion Model Predictions of the Joint Urban 2003 Field Experiment

S. Warner, Project Leader
J.F. Heagy
N. Platt
J.T. Urban

March 2007

IDA Paper P-4195
Log: H 07-000138

This work was conducted under contract DASW01-04-C-0003, Task DC-1-2607, for the Defense Threat Reduction Agency (DTRA). The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-4195

Comparisons of Transport and Dispersion Model Predictions of the Joint Urban 2003 Field Experiment

S. Warner, Project Leader
J.F. Heagy
N. Platt
J.T. Urban

PREFACE

This paper was prepared by the Institute for Defense Analyses (IDA) for the Defense Threat Reduction Agency (DTRA), in partial fulfillment of the task "Support for DTRA in the Validation Analysis of Hazardous Material Transport and Dispersion Prediction Models." The objective of this effort was to conduct analyses and special studies associated with the verification, validation, and accreditation (VV&A) of hazardous transport and dispersion prediction models. John R. Hannan and Richard N. Fry served as the DTRA project monitors for this task. The IDA Technical Review Committee was chaired by George E. Koleszar and consisted of Christine R. Bucher, Dennis F. DeRiggi, Vincent B. Lillard, Don A. Lloyd, and Davy Y. Lo.

The authors thank Richard J. Babarsky (National Ground Intelligence Center, Charlottesville, VA), Rong-Shyang Sheu (National Center for Atmospheric Research, Boulder, CO), and Ralph Gailis (Defense Science Technology Office, Melbourne, Australia) for useful discussions and access to various meteorological data assimilations and forecasts. The authors also acknowledge R. Ian Sykes and Doug S. Henn of Titan Corporation (Princeton, NJ), Cathy Dougherty and Thomas B. Harris of Science Applications International Corporation (King of Prussia, PA), Jacques Moussafir of ARIA Technologies (Boulogne-Billancourt**,** France), and John R. Hannan and Richard N. Fry of DTRA/Joint Science and Technology Office (Fort Belvoir, VA) for valuable discussions throughout this study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

The potential effects of the atmospheric release of hazardous materials continue to be of concern to the nation – a concern especially acute in more densely populated urban areas. In addition, military activities increasingly are conducted in urban, highly populated areas. Therefore, the possible release of hazardous materials in an urban environment is especially troubling to both civilian and military authorities.

In the case of hazardous material releases, effective mitigation in urban settings will require an understanding of the transport and dispersion of these hazards in the urban environment. The U.S. Departments of Defense (DoD) and Homeland Security (DHS) need to be able to estimate the effects of hazardous releases within an urban environment on the underlying population to aid planning, emergency response, and recovery efforts. These estimates require accurate knowledge of the concentrations of dispersed material in time and space. Especially desired are estimates of where and when relatively low-level human effects thresholds are exceeded.

Improved characterization and understanding of urban transport and dispersion (T&D) will allow for more robust modeling. The buildings located in urban regions, sometimes large and often closely packed, create their own roughness-induced boundary layers that can disperse toxic materials in ways that are not completely understood. Other features of the urban environment, including traffic-induced turbulence, heat island formation, flows associated with the deep street canyons of some large cities, the relative lack of moisture, and differential heating on building faces, also can have varied effects on the transport and dispersion of hazardous materials.

Until recently, urban transport and dispersion models could be divided into two main categories: (1) low-fidelity models (e.g., urban canopy models) that account for the large-scale effects of urban terrain, such as drag from buildings and boundary layer perturbations, and (2) high-fidelity models (e.g., computational fluid dynamics models) that include detailed representations of buildings, streets, and other urban features. "In between" lies a recent class of urban T&D models that take into account detailed urban features but employ empirical turbulence and wind profile parameterizations derived

from these urban details. Examples include the particle-based MESO/RUSTIC[1] model and the puff-based Hazard Prediction Assessment Capability (HPAC)[2] model. In this paper, we concentrate on the urbanized version of the HPAC model.

A few recent field experiments have included the release of environmentally safe, inert, tracer gases in urban environments. For example, tracer gases were released in Salt Lake City, UT, in 2000[3] and during the Mock Urban Setting Test (MUST) at Dugway Proving Ground.[4] An important use of the data collected during these field experiments is to provide support for the evaluation of the transport and dispersion models. Data collected during the 2000 Salt Lake City atmospheric tracer and meteorological study, referred to as Urban 2000, and MUST have been used to aid assessments of the validity of HPAC.[5, 6]

Under the joint sponsorship of the DoD (Defense Threat Reduction Agency - DTRA) and DHS, a series of tracer gas releases were carried out in Oklahoma City starting on 28 June and ending on 31 July 2003. This field experiment, referred to as "Joint Urban 2003," included ten intensive operating periods (IOPs), in which the tracer gas sulfur hexafluoride (SF_6) was released in downtown Oklahoma City.[7] In total,

---

1 Diehl, S. R., D. T. Smith, and M. Sydor, 1982: Random-walk simulation of gradient-transfer processes applied to dispersion of stack emission from coal-fired power plants. *J. Appl. Meteor.*, **21**, 69-83., and Burrows, D. A., R. Keith, S. Diehl and E. Hendricks, 2004: A fast running urban air flow model. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27. MESO is a Monte Carlo Lagrangian dispersion code and RUSTIC = Realistic Urban Spread and Transport of Intrusive Contaminants.

2 DTRA, 2001: The Hazard Prediction and Assessment Capability (HPAC) user's guide version 4.0.3. Prepared for the Defense Threat Reduction Agency by Science Applications International Corporation, Rep. HPAC-UGUIDE-02-U-RAC0, 602 pp.

3 Allwine, K. J., J. H. Shinn, G. E. Streit, K. L. Clawson, and M. Brown, 2002: Overview of URBAN 2000, A multiscale field study of dispersion through an urban environment. *Bull. Amer. Meteor. Soc.*, **83**, 521-536.

4 Biltoft, C. A., 2002: *Customer report for Mock Urban Setting Test,* DPG Document No. WDTC_FR-01-121, Meteorology and Obscurants Division, West Desert Test Center, U.S. Army Dugway Proving Ground, Dugway UT, 84022-5000, 58 pp.

5 Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of transport and dispersion model predictions of the URBAN 2000 field experiment, *J. Appl. Meteor*., **43**, 829-846.

6 Warner, S., N. Platt, J. F. Heagy, J. E. Jordan, and G. Bieberbach, 2006: Comparisons of transport and dispersion model predictions of the mock urban setting test field experiment. *J. Appl. Meteor. and Climatology,* **45**, 1414-1428.

7 Allwine, K. J., M. J. Leach, L. W. Stockham, J. S. Shinn, R. P. Hosker, J. F. Bowers, and J. C. Pace, 2004: Overview of joint urban 2003—An atmospheric dispersion study in Oklahoma City. *Symp. on Planning, Nowcasting and Forecasting in the Urban Zone,* Seattle, WA, Amer. Meteor. Soc., January 12-16.

twenty-nine 30-minute continuous SF_6 releases were accomplished with 2 hours of sampler monitoring following the start of each release.

This study examines the 29 continuous Joint Urban 2003 SF_6 releases and describes comparisons of the tracer gas sampler observations to Urban HPAC predictions. Several sets of predictions based on different configurations of Urban HPAC and varying meteorological inputs are examined in this paper. The preparation of these predictions – for example, the protocols required to create plausible predictions – is also described herein. In addition, comparisons between sets of Urban HPAC predictions are examined, and thus, relative performance among Urban HPAC configurations is assessed.

The results of this study are expected to provide information that supports two broad Government decision-making areas: (1) identification of promising urban transport and dispersion codes to be further developed and potentially integrated into other military hazardous assessment systems (e.g., Joint Effects Model) and (2) specification of best practices with respect to using HPAC in urban environments to include urban mode settings and appropriate and sufficient use of meteorological inputs. In addition to the above, the analyses and results of this report support the general further development of hazardous material assessment tools.

A. OVERVIEW OF JOINT URBAN 2003 (JU03)

As shown in Figure 1, four release locations, all in downtown Oklahoma City, were used for the 29 continuous SF_6 releases – referred to here as Park Avenue (PA), Botanical Gardens (BG), Westin-A (WA), and Westin-B (WB). The duration of each of these releases was 30 minutes. Of the 29 releases, 17 occurred during the day and 12 occurred at night.

This study focused on comparisons of predicted and observed concentrations near the surface. Figure 2 shows the typical locations of the SF_6 ground-based samplers[8] that were used for this study. These "ground-based" samplers were located at approximately 3 meters above ground level (AGL), typically mounted on light poles. Fifty-five samplers were located in the central business district (CBD), and 23 were located on an

[8] Chapter 1 provides additional descriptions of the samplers and data collected (e.g., 30-minute average SF_6 concentration).

arc located about 1 km from the release point. Similarly, 21 samplers were located on a 2-km arc and 21 were located on the 4-km arc.

Figure 1. Overhead View of Downtown Oklahoma City Showing the Four SF_6 Continuous Release Locations

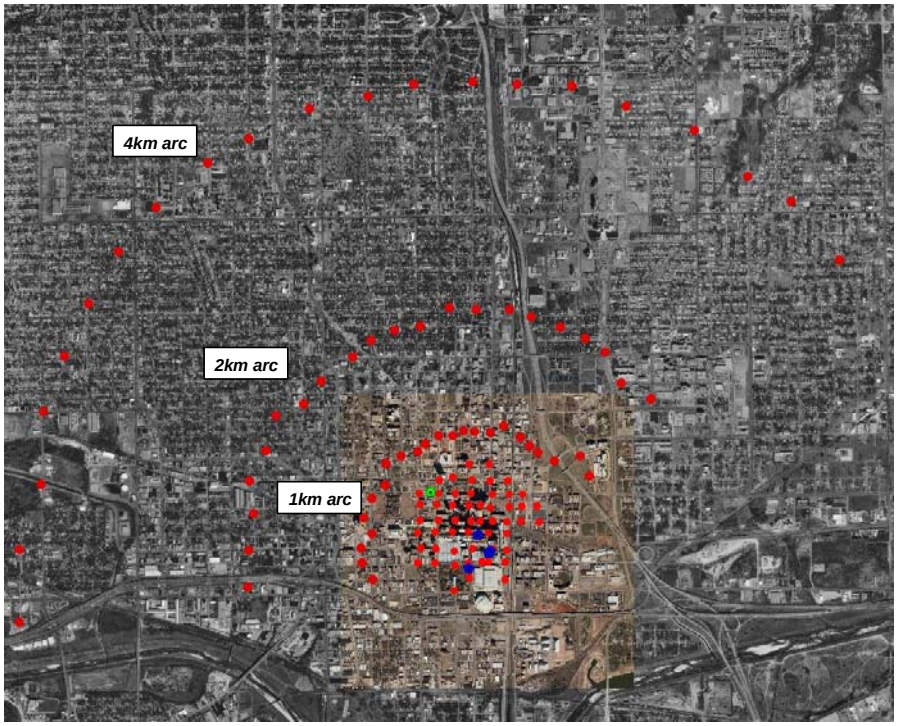


Figure 2. Locations of JU03 Surface Samplers

The blue pentagons correspond to the locations of the IOP releases (Figure 1) and the green pentagon corresponds to the location of a single mini-IOP that was designed to examine vertical dispersion at a single location.

Extensive meteorological measurements were made during the JU03 experiment. A brief description of some of the meteorological information used to create Urban HPAC predictions is described in the next section and Chapter 2 (and references therein) provide additional details and discussion.

B. CREATION OF URBAN HPAC PREDICTIONS OF JU03

The first goal of this study was to create plausible predictions of JU03 using the HPAC software.[9] A complete set of HPAC predictions of the JU03 field experiment have not previously been reported. For this study, we examined five types of Urban HPAC predictions: HPAC (v4.04 SP3) with surface type entered as "urban," denoted baseline or "UC" (for urban canopy); HPAC with the Urban Dispersion Model (UDM) toggled on, denoted "DM"; HPAC with the Urban Windfield Module (UWM) toggled on, denoted "WM"; and HPAC with both the UDM and the UWM toggled on, denoted "DW." We also examined predictions created with the newest urban feature, Micro SWIFT SPRAY (MSS), denoted as "MS." In general, default settings were used to create HPAC predictions.

Given observations or predictions generated by a mesoscale meteorological model, HPAC can create mass-consistent wind fields that can be used to transport the hazardous material. Within HPAC two weather modules can be used to prepare these mass-consistent wind fields: Stationary Wind Fit and Turbulence (SWIFT) and a Mass-Consistent SCIPUFF (MC-SCIPUFF) algorithm. In the case of SWIFT, the underlying topography is required. It should be noted that neither SWIFT nor MC-SCIPUFF account explicitly for wind speed profiles below the mean building height (i.e., within the urban canopy).

1. Urban HPAC Transport and Dispersion Modes

For urban applications of HPAC, the vertical wind and turbulence profiles can be modified to account for urban effects. This mode of operation is referred to as "UC" and is considered as a baseline for comparison in this study (and further described in Chapter 1). Brief overviews of the Urban HPAC modules UDM, UWM, and MSS are described below and more detail is provided in Chapter 1 and references therein.

---

[9] For hazardous material transport and dispersion, HPAC uses the Second-Order Closure Integrated Puff (SCIPUFF) model and an associated mean wind field model. (See Chapter 1 for more details.)

- **UDM:** The UDM is designed to compute the transport and dispersion of an instantaneous discharge ("a puff" or train(s) of puffs) of pollutant released over a surface containing a mixture of open and urban areas. UDM considers variations in dispersion rates as a function of surface changes and direct interaction of pollutant cloud with surface obstacles. UDM is based on ensemble mean Gaussian puff dispersion methodology but allows surface obstacles to modify the dispersion patterns. UDM empirical parameters are set based on extensive wind tunnel experiments. Three regimes for calculation are identified within UDM: open regime, urban regime, and longer-range regime. If a single obstacle is encountered in the open regime, the puff will interact and be partitioned into entrained and unentrained fractions. The unentrained fraction will continue with some enhanced dispersion due to increased turbulence in the recovery region. The entrained fraction remains in a recirculating flow in the obstacle's wake, with a modeled characteristic residence time. As the entrained flow escapes, it is transported and dispersed separately from the unentrained fraction.
- **UWM:** The UWM predicts steady-state winds (speed and direction) inside the urban boundary layer using a canopy parameterization. UWM-generated average winds can then be used by Urban HPAC (for example, with or without UDM toggled on) to drive material transport and dispersion. UWM is a computational fluid dynamics (CFD) code that is designed to provide a computationally fast solution, within a CFD framework, by considering spatially averaged obstacle effects. Therefore, the predicted winds of UWM represent spatio-temporal averages – spatial averages of the Reynolds-averaged equations of motion – which allows for fast solutions. This spatial averaging causes the momentum and energy equations to have extra terms that correspond to important physical processes, which include drag forces and turbulent energy production and are parameterized within UWM's canopy approach. Initial conditions for the UWM within HPAC can be set by including a mass-consistent, three-dimensional, gridded wind field based on observations (e.g., SWIFT or MC-SCIPUFF) or by providing a gridded numerical weather prediction.
- **MSS:** MSS is designed to provide a fast computation of the wind field within the urban environment while accounting for an exact representation of the buildings, e.g., as generated by Geographic Information Systems. Within HPAC, SWIFT or MC-SCIPUFF can provide initial boundary meteorological conditions. Given information of the local buildings (locations, shapes, and sizes), Micro SWIFT

creates a modified wind field by creating zones where the flow is modified according to the buildings' locations, and flow is adjusted to satisfy the continuity equation and impermeability on the ground and on the buildings' walls. Micro SWIFT also derives a diagnostic turbulence – diffusive coefficients and the turbulent kinetic energy dissipation rate – by considering the distance to the nearest obstacle as a mixing length and using this value for wind field local shear. Micro SPRAY is a Lagrangian particle dispersion model that can account for obstacles. Dispersion is simulated by following the movement of a large number of fictitious particles, each representing a portion of the original released mass. The motions of the particles are obtained by applying an equation of motion that has two components – a mean component, which follows the local winds as defined by Micro SWIFT, and a stochastic component. The stochastic component of the particle motion follows a scheme that includes a stochastic Gaussian term.

2. Meteorological Input Options Used for Urban HPAC Mode Comparisons

A large variety of meteorological measurements were collected during JU03. The main goal of this paper is to describe the comparative results of JU03 predictions created with varying Urban HPAC configurations – UC, WM, DM, DW, and MS. Several meteorological ("MET") input options were examined in this study, in part, to understand how relative Urban HPAC mode performance varies for different reasonable MET options. This section provides a brief description of the five MET input options that were selected for this comparative study.

The shorthand notations for the five MET options that were chosen for this study are (1) BAS, (2) BRB, (3) PNA, (4) ACA, and (5) PO7. Each MET option is briefly described below. Additional details associated with the chosen meteorological input options are provided in Chapter 2.

- **BAS:** The BAS MET option was designed to correspond to a baseline situation where the meteorological information is consistent with what could have been retrieved from the DTRA meteorological server at some point (~ 2 hours or more) after the release. That is, this information corresponds to assimilated observations during the release as opposed to forecast information. Meteorological sources within 30 km of Oklahoma City were considered for BAS. Surface wind velocity observations from four stations between 12 and 28 km from the downtown area and upper air wind velocity observations from a station 28 km southeast of

Oklahoma City were used. The diagnostic wind field model, SWIFT, was used to create gridded wind fields from the BAS input meteorological information.

- **BRB:** The BRB MET option was designed to correspond to a baseline situation in which a gridded numerical weather assimilation was used as input to Urban HPAC. The BRB MET option corresponded to a Global Climatological Analysis Tool (GCAT) prediction of wind velocity profiles (i.e., wind velocity as a function of height AGL) at many grid locations. These files can be thought of as surrogates for "gridded" numerical weather assimilations that could be available on the DTRA meteorological server several hours after an event. SWIFT was used to create gridded wind fields from the BRB input meteorological information
- **PNA:** The PNA MET option corresponded to using both the SODAR[10] and vertical profiler observations that were available at an upwind (~ 1.6 km) site. The SODAR and profiler provide wind speeds and directions as a function of altitude at a single geographic location – a "vertical profile." Previous studies[11] of Urban HPAC suggested that a single measured upwind vertical profile can represent a satisfactory input to create reasonable urban predictions. Therefore, the PNA MET option allows for the further testing of this hypothesis. For the PNA MET option, MC-SCIPUFF was used to create gridded wind fields instead of SWIFT.[12]
- **ACA:** The ACA MET option corresponded to using both the SODAR and profiler observations that were available at the downwind site (~ 4 km). This MET option was considered particularly useful for comparison and contrast with PNA – the single upwind site option. For the ACA MET option, MC-SCIPUFF was used to create gridded wind fields.[13]
- **PO7:** The PO7 MET option corresponded to a set of observations from a single location 40 meters AGL on the roof of the Oklahoma City Post Office building (just upwind of the downtown). This option corresponds to using a single downtown observation as input for the Urban HPAC predictions. During a

---

10 SODAR = Sonic Detection and Ranging or simply Acoustic Sounder.

11 See references cited in footnotes 5 and 6.

12 A SWIFT software error, that caused an HPAC abort, was encountered for some releases when using the PNA MET option. Therefore, MC-SCIPUFF was used for all PNA-based predictions.

13 As was the case for the PNA MET option, MC-SCIPUFF was used with the ACA option because SWIFT errors were occasionally encountered.

previous study of urban atmospheric transport and dispersion ("Urban 2000" in Salt Lake City), it was found that using a single downtown building top measurement resulted in relatively worse predictions in terms of fits to the observations when compared with the other MET options that were examined. Therefore, the PO7 MET option allows for the reconsideration of this concept, i.e., using a set of observations from a single building top as input for hazardous material transport and dispersion predictions, albeit this time a somewhat upwind building as opposed to a downtown building.

3. Summary of *Compared* Sets of JU03 Predictions

Twenty-five sets of Urban HPAC predictions were generated as described in Table 1. Detailed descriptions of the specific protocols followed to create these JU03 predictions are described in Chapter 2.

Table 1. Shorthand Notations for the 25 Model Combinations That Were Examined

	Transport and Dispersion Model Mode				
MET Options	**UC**	**WM**	**DM**	**DW**	**MS**
BAS	BAS_UC	BAS_WM	BAS_DM	BAS_DW	BAS_MS
BRB	BRB_UC	BRB_WM	BRB_DM	BRB_DW	BRB_MS
PNA	PNA_UC	PNA_WM	PNA_DM	PNA_DW	PNA_MS
ACA	ACA_UC	ACA_WM	ACA_DM	ACA_DW	ACA_MS
PO7	PO7_UC	PO7_WM	PO7_DM	PO7_DW	PO7_MS

The methodologies used for the comparisons of this study have been previously described[14] and are further discussed in Chapter 2. A variety of statistical metrics to examine bias, scatter, and correlation were examined as well as a user-oriented measure of effectiveness (MOE)[15] that allowed for assessments of the ability of the model to predict the "hazardous" region (i.e., region above a concentration threshold of interest). Procedures for detecting statistically significant differences between metrics and

---

14 See the reference cited in footnote 5.

15 Warner, S., N. Platt, and J. F. Heagy, 2004: User-oriented two-dimensional measure of effectiveness for the evaluation of transport and dispersion Models, *J. Appl. Meteor*., **43**, 58-73.

estimating confidence intervals associated with point estimates of metrics are also described in Chapter 2. In addition, and as important as any metric, comparative plots of model predictions and observations were created and scrutinized for all releases.

For each comparison, 30-minute average concentration comparisons are examined in this report in the CBD and separately for all of the arc-based samplers (1, 2, and 4 km arcs). It was recognized early in this analysis that model performance varied greatly as a function of the release time – day or night. Therefore, analyses were done separately for the day and night releases. The focus of the comparisons discussed here is on the following metrics: Normalized Absolute Difference (NAD) and Normalized Mean Square Error (NMSE), both measures of scatter between observations and predictions, Fractional Bias (FB), a measure of bias (e.g., over- or under-prediction), and the aforementioned MOE. As a part of these studies, several other metrics (e.g., fraction of the predictions within a factor of 2, 5, or 10 of the predictions – FAC2, FAC5, and FAC10) and conditions (e.g., time resolution – 2 hour average; time period – 1st 30-minute period; location – 1 km arc only) were examined and are identified in Chapter 2. In total, over 80,000 metrics were computed for different model configurations and conditions.

The next two sections of this Summary describe the principal findings and conclusions of this study.

C. COMPARISONS OF PREDICTIONS AND JU03 OBSERVATIONS

1. Summary of Urban HPAC Mode Comparisons

Table 2 shows, for each of the five MET input options that were considered, the Urban HPAC modes that resulted in the least scatter, i.e., the best fit to the observations. When applied to observations and predictions paired in space and time, scatter-based metrics allow for the evaluation of how well the model predicted the location and timing (at least for 30-minute averages examined here) of the observations. For this reason, we consider the three scatter-evaluating metrics discussed in this report – NAD, NMSE, and the concentration-based MOE – as particularly important measures of model predictive performance.

Table 2 identifies the Urban HPAC modes that resulted in relative (and statistically significant) improvement for the five MET input options and the four conditions (day and night, CBD and arcs) that were examined. For example, the "(**MS**,**DM**,**DW**)/(**UC**,**WM**)" for the BAS-Night, CBD, cell in Table 2, implies that for

two of the three scatter-related metrics, **MS**, **DM**, **DW** outperformed (i.e., "less scatter" and statistically significant for two of the three scatter-related metrics) **UC** and **WM**. Several robust conclusions are apparent from Table 2 and additional analyses are discussed in Chapter 3.

Table 2. Urban HPAC Modes, for Five MET Input Options, That Led to Improved Predictive Performance of JU03 Releases Based on Measures of Predicted/Observed Scatter (Concentration-Based MOE, NAD, and NMSE)[a]

Condition	BAS (SWIFT)	BRB (SWIFT)	PO7 (SWIFT)	PNA (MC-SCIPUFF)	ACA (MC-SCIPUFF)
Day CBD	DW/DM	mixed	mixed	(UC,WM,DM,DW) /MS and DW/DM	DW/MS
Day Arcs	(MS,DW) /(UC,WM)	mixed	mixed	mixed	no differences
Night CBD	(MS,DM,DW) /(UC,WM)	(MS,DM,DW) /(UC,WM)	(MS,DM,DW) /(UC,WM) and MS / (DM,DW)	mixed	no differences
Night Arcs	(MS,DM,DW) /(UC,WM) and DM/DW	(MS,DM,DW) /(UC,WM)	(MS,DM,DW) /(UC,WM) and MS/ (DM,DW)	mixed	MS / (UC,WM,DM)

[a] The XX/YY nomenclature denotes model(s) XX had statistically significant relative improvement over model(s) YY. For this table, the designation implies that, for at least two of the three scatter-related metrics (concentration-based MOE, NAD, and NMSE), XX showed a statistically significant improvement relative to YY. The word "mixed" implies that there was not a consistent finding of one model or models over others.

a. Day vs. Night Releases and Predictions

First, there was a substantial difference in the performance of Urban HPAC as a function of day and night. Figure 3 shows comparative NAD results for the five MET options and five Urban HPAC modes for both day and night and for the CBD samplers. For the SWIFT-associated MET options – BAS, BRB, and PO7 – Urban HPAC predictions resulted in substantially more scatter at night than during the day, with the exception of **MS**. For the MC-SCIPUFF-associated MET input options, the scatter results were much more similar for the day and night Urban HPAC predictions, with perhaps some evidence of improved performance during the day for PNA and ACA. Examinations of arc-based results showed similar behavior to that described above for the CBD.

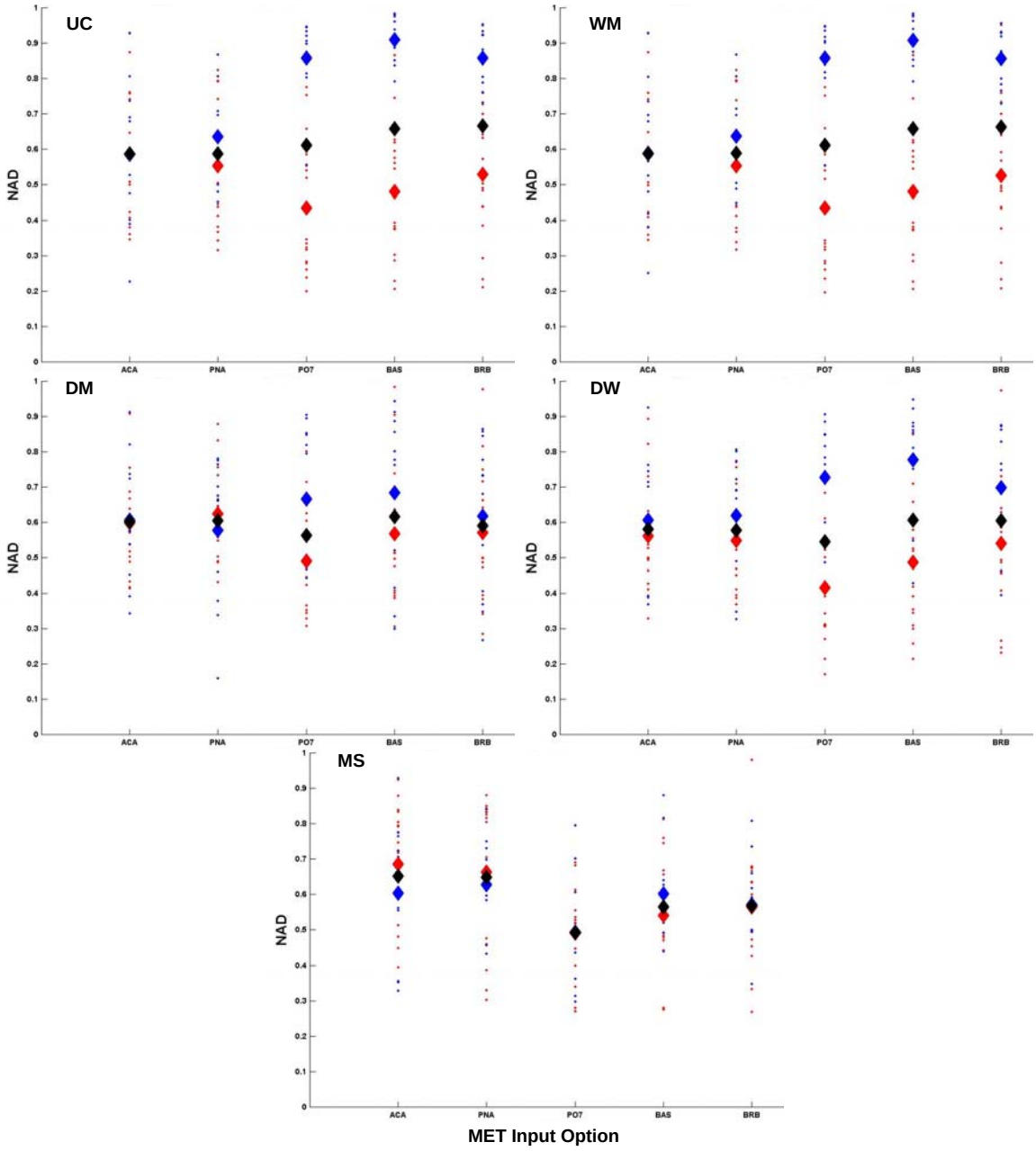


Figure 3. Comparisons of NAD Values for Urban HPAC Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 Within the CBD Using the Five MET Input Options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB)

The smaller colored points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD values (day = red and night = blue) with the large black diamond representing the overall average for all 29 releases.

For all five MET options, daytime releases tended to be under-predicted (30-minute average concentrations at the surface samplers in the CBD and on the arcs) and nighttime releases tended to be over-predicted. Figure 4 compares day and night FB

values for the CBD samplers and for the 25 sets of predictions. Examinations of arc-based results showed similar behavior to that described above for the CBD.

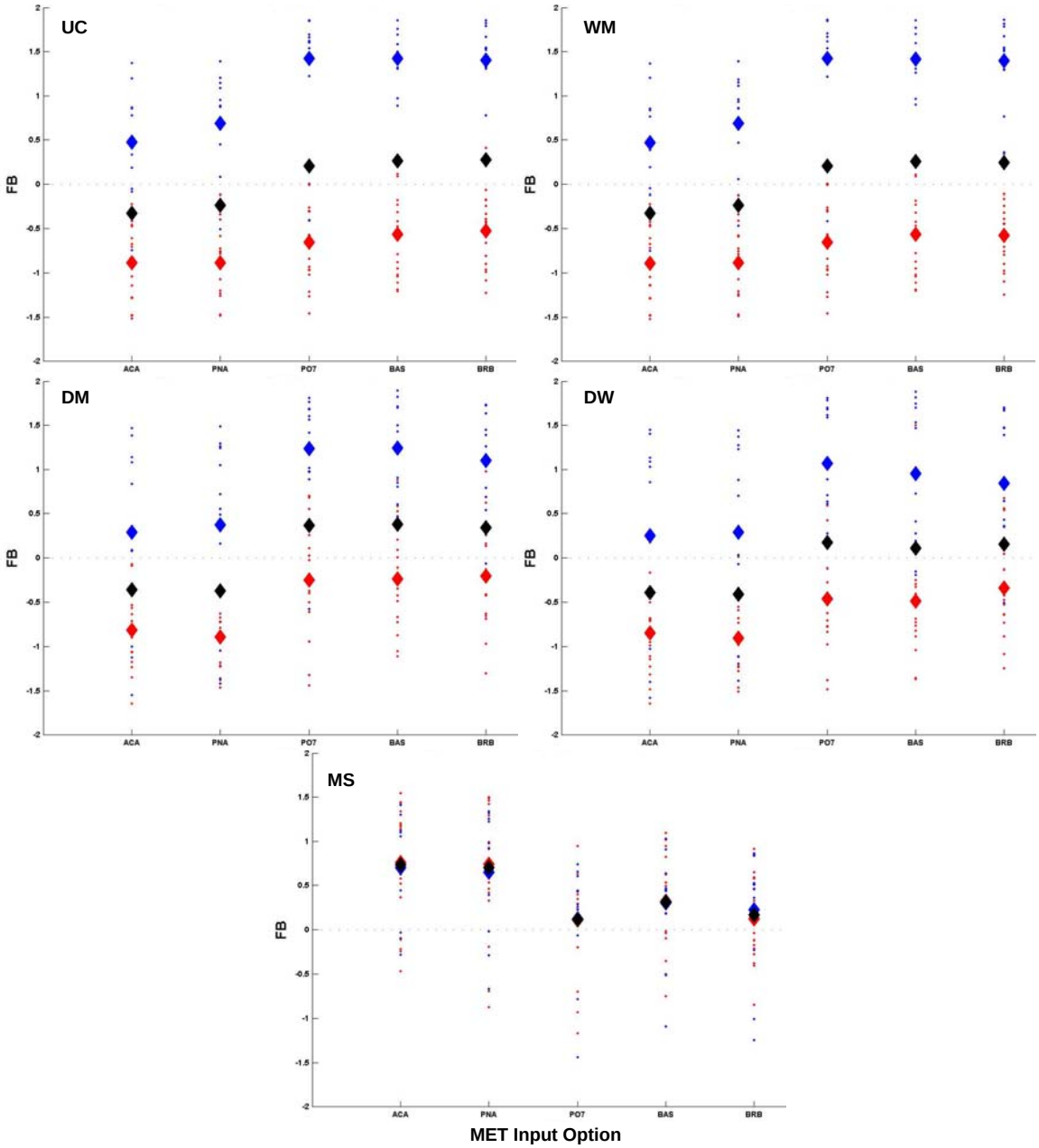


Figure 4. Comparisons of FB Values for Urban HPAC Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 Within the CBD Using the Five MET Input Options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB)

The smaller colored points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB values (day = red and night = blue) with the large black diamond representing the overall average for all 29 releases.

b. MSS Model Performance Behavior Differs From Other Urban HPAC Modes

With respect to the under- and over-predictions described above, the **MS** mode typically led to less under-prediction during the day *and* less over-prediction at night than the other Urban HPAC modes (there were some minor exceptions where **DM** and **DW** modes were similar to **MS**). Typically, the **MS** mode resulted in the least biased predictions of the 30-minute average concentrations at the surface samplers (CBD and arcs).

Tables 3 and 4 list the NAD and NMSE values, respectively, from least scatter (best) to most scatter (worst). **MS**-based predictions resulted in the least scatter in seven of the eight categories (day/night, CBD/arcs, and NAD/NMSE). The PO7_**DW** is the sole exception to the above, having the best NAD value for the day-CBD condition.

Table 3. NAD Values for 25 Sets of Urban HPAC Predictions of JU03 Ordered from Least to Most Scatter. Values are shown for day and night and for CBD and Arcs.

	Day	CBD	Metric	Night	CBD	Metric	Day	Arcs	Metric	Night	Arcs	Metric
Rank	Mode	MET	NAD	Mode	MET	NAD	Mode	MET	NAD	Mode	MET	NAD
1	DW	PO7	0.42	MS	PO7	0.49	MS	PNA	0.31	MS	ACA	0.35
2	WM	PO7	0.44	MS	BRB	0.57	DW	PNA	0.32	DW	ACA	0.39
3	UC	PO7	0.44	DM	PNA	0.58	DW	PO7	0.34	WM	ACA	0.40
4	WM	BAS	0.48	UC	ACA	0.59	WM	PNA	0.34	UC	ACA	0.40
5	UC	BAS	0.48	WM	ACA	0.59	UC	PNA	0.34	DM	ACA	0.41
6	DW	BAS	0.49	MS	BAS	0.60	UC	PO7	0.35	DM	PNA	0.45
7	DM	PO7	0.49	MS	ACA	0.60	WM	PO7	0.35	DW	PNA	0.47
8	MS	PO7	0.49	DM	ACA	0.61	DM	PO7	0.35	MS	PNA	0.48
9	WM	BRB	0.53	DW	ACA	0.61	MS	PO7	0.36	MS	PO7	0.50
10	UC	BRB	0.53	DM	BRB	0.62	DM	PNA	0.36	UC	PNA	0.51
11	MS	BAS	0.54	DW	PNA	0.62	MS	BAS	0.37	WM	PNA	0.51
12	DW	BRB	0.54	MS	PNA	0.63	DW	ACA	0.39	DM	BRB	0.53
13	DW	PNA	0.55	UC	PNA	0.64	MS	ACA	0.40	MS	BRB	0.57
14	WM	PNA	0.55	WM	PNA	0.64	DM	ACA	0.41	DW	BRB	0.58
15	UC	PNA	0.55	DM	PO7	0.67	UC	ACA	0.41	DM	BAS	0.60
16	DW	ACA	0.56	DM	BAS	0.68	DW	BAS	0.41	DM	PO7	0.63
17	MS	BRB	0.57	DW	BRB	0.70	WM	ACA	0.41	MS	BAS	0.64
18	DM	BAS	0.57	DW	PO7	0.73	WM	BRB	0.43	DW	PO7	0.64
19	DM	BRB	0.57	DW	BAS	0.78	WM	BAS	0.44	DW	BAS	0.69
20	UC	ACA	0.59	WM	BRB	0.86	UC	BAS	0.44	UC	BRB	0.75
21	WM	ACA	0.59	UC	BRB	0.86	UC	BRB	0.44	WM	BRB	0.75
22	DM	ACA	0.60	WM	PO7	0.86	DW	BRB	0.45	WM	PO7	0.77
23	DM	PNA	0.62	UC	PO7	0.86	DM	BAS	0.46	UC	PO7	0.77
24	MS	PNA	0.66	WM	BAS	0.91	DM	BRB	0.47	WM	BAS	0.86
25	MS	ACA	0.69	UC	BAS	0.91	MS	BRB	0.57	UC	BAS	0.86

While it typically took less than 5 minutes to generate a **UC** or **DM** mode prediction of a single release (tracking the plume for 2 hours), **MS** predictions, at least at the resolutions used for the predictions that we created, took, on average, about 60 minutes per release. We also created a lower resolution set of **MS** predictions, for which each prediction took about 30 minutes and found substantially similar results to those reported here for the higher resolution runs. With the version of MSS that we had access to, we could not create even lower resolution **MS** predictions.

Table 4. NMSE Values for 25 Sets of Urban HPAC Predictions of JU03 Ordered from Least to Most Scatter. Values are shown for day and night and for CBD and Arcs.

	Day	CBD	Metric	Night	CBD	Metric	Day	Arcs	Metric	Night	Arcs	Metric
Rank	Mode	MET	NMSE	Mode	MET	NMSE	Mode	MET	NMSE	Mode	MET	NMSE
1	MS	PO7	16.91	MS	PO7	27.68	MS	PNA	7.69	MS	ACA	9.49
2	DW	BAS	17.68	MS	BAS	28.58	MS	PO7	9.34	MS	PNA	11.58
3	MS	BRB	17.94	MS	BRB	33.38	DW	PNA	9.40	DW	ACA	12.27
4	DW	PO7	20.08	WM	PNA	59.63	MS	BAS	10.29	DM	PNA	12.47
5	DW	BRB	20.61	UC	PNA	60.41	DM	PO7	10.47	DM	ACA	13.26
6	DM	BRB	21.08	MS	PNA	63.36	DW	PO7	11.27	DW	PNA	14.09
7	DM	PO7	22.70	WM	ACA	71.02	WM	PNA	11.41	UC	PNA	14.43
8	MS	BAS	24.37	UC	ACA	71.10	UC	PNA	11.54	WM	ACA	14.47
9	WM	BAS	27.15	MS	ACA	77.88	UC	PO7	11.76	WM	PNA	14.65
10	UC	BAS	27.22	DM	PNA	78.59	DM	PNA	11.80	UC	ACA	14.71
11	UC	BRB	29.07	DW	ACA	90.38	WM	PO7	11.80	MS	PO7	14.86
12	WM	PO7	29.27	DW	PNA	91.96	DM	BRB	11.85	MS	BAS	18.32
13	UC	PO7	29.33	DM	ACA	96.66	DW	BRB	12.11	DM	BAS	24.96
14	WM	BRB	29.66	DM	BRB	133.79	DM	ACA	12.88	DM	BRB	26.17
15	DM	BAS	31.48	DM	PO7	143.00	DW	BAS	13.07	MS	BRB	33.38
16	DW	PNA	33.17	DW	BRB	150.00	DM	BAS	13.42	DW	BRB	34.99
17	DW	ACA	33.97	DW	PO7	159.87	UC	BRB	13.45	DM	PO7	35.54
18	DM	ACA	38.54	DM	BAS	185.90	WM	BRB	13.62	DW	PO7	49.52
19	DM	PNA	38.81	DW	BAS	218.92	DW	ACA	15.30	DW	BAS	50.34
20	WM	PNA	45.09	WM	BAS	340.29	MS	ACA	15.57	UC	BRB	120.48
21	UC	PNA	45.22	UC	BAS	346.48	UC	ACA	17.04	WM	BRB	121.84
22	UC	ACA	50.18	UC	BRB	488.84	WM	ACA	17.13	WM	PO7	141.98
23	WM	ACA	50.36	WM	BRB	493.54	MS	BRB	17.94	UC	PO7	143.53
24	MS	ACA	100.40	UC	PO7	506.26	WM	BAS	19.17	WM	BAS	161.63
25	MS	PNA	119.28	WM	PO7	521.59	UC	BAS	19.18	UC	BAS	162.43

c. Relative Urban HPAC Mode Performance for Nighttime Releases: MS, DM, and DW Represented Improvements

An additional important result is that for the nighttime releases, the **MS**, **DM**, and **DW** modes offer improvement over the **UC** and **WM** modes for the three MET input options that invoked SWIFT. This finding was true for the samplers in the CBD and for the samplers along the arcs. This result can be considered especially important because the use of SWIFT corresponds to a recommended and default mode of Urban HPAC. In addition, these MET options, particularly BAS and BRB, appear to correspond to reasonably realistic and potential operational applications of Urban HPAC. We also found that adding UWM to UDM to create the **DW** mode did not lead to substantial or consistent significant improvements relative to using UDM alone, i.e., **DM**. This result is entirely consistent with past studies of the Urban 2000 and MUST field trials. It also should be noted that the **DW** predictions (as we ran them) took approximately 80 minutes longer per release than the corresponding **DM** prediction. These results, and past findings, call into question the value of including UWM, at least as we have been able to implement this feature.

For the nighttime releases and the MC-SCIPUFF-associated MET input options – PNA and ACA – results were mixed with no Urban HPAC mode consistently offering

improvement, although the **MS** mode did so for the ACA MET option (at least relative to **UC**, **WM**, and **DM**).

d. Relative Urban HPAC Mode Performance for Daytime Releases Was Mixed and Inconsistent

For the daytime releases, no consistent trend was found. For example, for the BAS-associated predictions on the arcs, the **MS**, **DM**, and **DW** modes offer improvement (e.g., less scatter) over the **UC** and **WM** modes, but for the PNA-associated predictions in the CBD, the **UC**, **WM**, and **DW** resulted in improved scatter relative to the **MS** and **DM** modes. However, in the latter PNA-based case, the observed improvements in scatter for the **UC**, **WM**, and **DW** predictions come at the cost of a large under-prediction relative to **MS**.

2. Concentration-Based Versus Threshold-Based MOE Values[16]

Predictions of exceeding a relatively low concentration threshold (5 × and 50 × background) were substantially more accurate (as measured by the MOE) than predictions of absolute 30-minute average concentrations. Figure 5 shows an example comparison of the concentration-based and threshold-based MOE values for the BAS_MS predictions. MOE point estimates for the 17 daytime and 12 nighttime releases are shown, and substantial "movement" of the MOE values toward the perfect value of (1,1) is shown. These results, and similar ones found for the other Urban HPAC mode/MET option combinations, indicate substantial improvements when the thresholds are examined. This result is entirely consistent with past studies of the Urban 2000 and MUST field trials. An important implication of the above finding is that using Urban HPAC to predict the extent (in time and space) to which a relatively low threshold is (or might be) exceeded is likely to lead to a more accurate representation of a hazardous release (or area) than using Urban HPAC to predict the actual concentrations in time and space (e.g., perhaps needed for a detailed and complete assessment of expected casualties given a human effects model that requires concentration-time histories as a function of location).[17]

---

16 A detailed description of the atmospheric transport and dispersion two-dimensional user-oriented MOE value and its application can be found in Chapter 2 and references therein.

17 A complete set of figures that compare MOE point estimates for all 25 sets of Urban HPAC predictions for the CBD and on the arcs, during the day and at night, and for the concentration-based, 25 parts per trillion (ppt) threshold-based, and 250 ppt threshold-based calculations is provided in Appendix B.

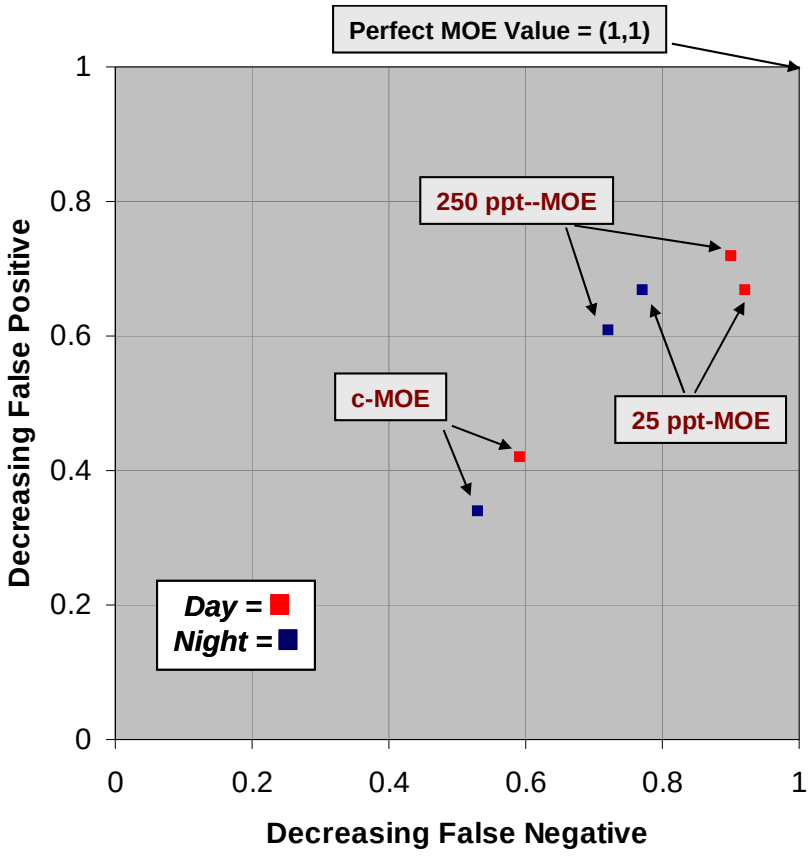


Figure 5. Comparisons of Concentration-Based MOE (c-MOE) and Threshold-Based MOE (25 ppt- and 250 ppt-MOE) Values for BAS_MS Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 within the CBD

3. Brief Comparison on MET Input Options

As has been previously discussed, day releases were generally under-predicted and the night releases were generally over-predicted both within the CBD and on the arcs. The **MS** mode represented an exception to the above, particularly for the CBD samplers.

A substantial difference between the two MC-SCIPUFF-associated MET options – PNA and ACA – and the SWIFT-associated MET options – BAS, BRB, and PO7 – was revealed. For the **UC**, **WM**, **DM**, and **DW** modes, the PNA and ACA options resulted in less material being predicted at the surface samplers, e.g., less over-prediction at night and more under-prediction during the day relative to the other MET options. For example, examination of Tables 3 and 4 show that for the nighttime releases on the arcs, the best five (of 25) NAD and NMSE values, respectively, were associated with the MC-

SCIPUFF-based predictions. Furthermore, the SWIFT-based predictions for the **UC**, **WM**, **DM**, and **DW** modes all led to substantial over-predictions (e.g., see Figure 3-52). This type of behavior could be caused by a faster wind speed being associated with the MC-SCIPUFF options relative to the SWIFT options. In this case, the faster wind speed simply blows material past the samplers too quickly and results in less material predicted in the 30-minute averages. Then, the faster winds reduce the over-prediction and improve statistical performance. This faster wind speed hypothesis was supported by comparisons of contour plots associated with differing predictions and observations. At this point, one interpretation of this result is that it mainly reflects compensating errors – substantial over-prediction and too fast winds. Another interpretation would be that some other combination of problems, perhaps including SWIFT-based winds that are too slow, could be the ultimate cause. Considerations of fundamental differences between SWIFT and MC-SCIPUFF as well as how each model component is implemented within HPAC and how such differences may account for the above findings, is an ongoing research topic in our group.

D. COMPARISON OF JU03 (OKLAHOMA CITY) AND PREVIOUS FINDINGS: URBAN 2000 (SALT LAKE CITY) AND MUST

The releases associated with the Urban 2000 and MUST field experiments were conducted at night or in the very early morning while atmospheric conditions were still relatively stable. Since there were substantial differences in Urban HPAC model behavior and performance as a function of day and night, the most appropriate JU03 results for comparison to Urban 2000 and MUST are those associated with the nighttime releases. Similarly, previous studies included the **UC**, **WM**, **DM**, and **DW** modes but not the **MS** mode, as it was not available at those times. Therefore, for these comparisons to previous studies we do not consider the more recent **MS** JU03 predictions. The previously created Urban 2000 and MUST predictions included the use of SWIFT in all cases. In fact, releases that led to SWIFT-generated errors were not considered in these previous studies, that is, no MC-SCIPUFF-based predictions were examined in these previous studies. Given the above considerations, JU03 findings that are particularly consistent with previous Urban 2000 and MUST conclusions include:

- In general, Urban HPAC modes led to over-predictions of the surface sampler concentrations. For JU03, the median nighttime FB value (not considering the **MS** mode) for the CBD and the arcs was 1.01 and 0.79 (or over-prediction factors of about 3.0 and 2.3), respectively. For Urban 2000, the comparable FB values

(for the CBD and on the arcs, respectively) were 0.48 (over-prediction factor of about 1.6) and 0.70 (over-prediction factor of about 2.1). For MUST, over-prediction factors, across the entire 200-meters square of samplers, varied by Urban HPAC mode from about 1.0 to 3.0.

- We also found for all three field experiments (JU03, Urban 2000, and MUST) that adding UWM to UDM to create the **DW** mode did not lead to *substantial or consistent* significant improvements relative to using UDM alone, i.e., **DM**.
- For all three field trial studies, predictions of exceeding a relatively low concentration threshold were more accurate (as measured by the MOE) than predictions of absolute 30-minute average concentrations.

With respect to NAD, the JU03 Urban HPAC predictions generally resulted in less scatter than the comparable Urban 2000 predictions. For example for Urban 2000, the downtown and arc-based NAD values were never less than 0.60 for any of the 20 sets of predictions that were examined and when considering 30-minute average concentrations. A few sets of Urban HPAC JU03 nighttime predictions (after excluding **MS** for these comparisons) resulted in NAD values less than 0.60, including SWIFT-based values on the arcs at night of 0.53 and 0.58 for the BRB-based **DM** and **DW** predictions, respectively. For MUST, NAD values are not easily comparable as the time resolutions that were examined were 10s, 60s, 300s, about 15 minutes are substantially different from those examined during JU03 and Urban 2000. Nonetheless, for the ≈15 minute time resolution, NAD values for the 20 sets of Urban HPAC MUST predictions were never less than 0.45, and typically above about 0.55. The relative improvement in model fit (less scatter as measured by NAD) for the JU03 predictions relative to those of Urban 2000 could be partially explained by improved Urban HPAC models available since the time of the Urban 2000 study *and/or* improved MET inputs used in this JU03 study, i.e., MET inputs that better represent the actual winds that transport the plume.

The FAC2 metric considers the ratios between the predictions and the observations at each point in space and time – here for 30-minute average concentrations. This metric is equally sensitive to the smaller and larger concentrations, whereas metrics such as NAD, NMSE, and the concentration-based MOE can be dominated by the larger concentrations. As such, FAC2 can be particularly sensitive to mismatches in plume transport direction (e.g., several samplers with relatively small observed concentrations missed on one side of the plume because of a 10-20 degree transport error could easily lead to predictions and observations differing by more than a factor of two). This metric

can also be very sensitive to the underlying sampling distribution and the data processing protocols. For comparisons between different field experiments, we followed consistent protocols in order to allow for fair comparisons between field trials.[18] Table 5 compares the range of FAC2 values for the comparable JU03 and Urban 2000 predictions. Table 5 suggests a substantial improvement associated with the JU03 predictions, perhaps due to improved MET inputs.

It was postulated during the analysis of the Urban 2000 field experiment and Urban HPAC predictions that the terrain associated with Salt Lake City represented a challenging and important feature that could influence the wind fields substantially. For example, a mesoscale numerical forecast that was used as MET input for the Urban 2000 study ("OMEGA") led to some of the best Urban 2000 predictions yet was recognized as missing the plume direction on the arcs.

We also examined a set of JU03 Urban HPAC predictions that used wind observations from the Botanical Gardens (Figure 1) in downtown Oklahoma City. Our analysis of the predictions that resulted from using this Botanical Gardens ("BGS") MET input option suggested that the wind directions were not well matched to the actual directions relative to the other MET options that were examined. This MET option was rejected for the final comparative analyses because it was determined to sometimes miss the wind directions and speeds significantly (Chapter 2). The FAC2 values for the BGS-based Urban HPAC predictions of JU03 are shown in the last column of Table 5.

Table 5. Range [Mean and Median], Across Modes and MET Input Options, of FAC2 Values for Urban HPAC Predictions of JU03 and Urban 2000 (MS-based JU03 predictions are not considered for these comparisons.)

Condition / Field Trial	JU03	Urban 2000	JU03 using BGS MET Input Option
Day	0.44-0.56 [0.49/0.49]	Not available	0.44-0.46 [0.45/0.45]
Night	0.35-0.51 [0.42/0.40]	0.18-0.26 [0.23/0.23]	0.18-0.25 [0.22/0.21]

The nighttime FAC2 values associated with the BGS predictions are quite similar to those previously reported for Urban 2000. The suggestion here is that improved MET

[18] Additional description of the chosen data protocol and sensitivity analyses associated with FAC2 computations and data protocol decisions are discussed in Chapter 3.

input options that better represented the actual winds were available and used for the JU03 predictions, especially relative to the comparable 30-minute average concentration Urban 2000 predictions. Future efforts are planned to evaluate the latest version of Urban HPAC (including MSS) using the Urban 2000 field experiment. Such efforts may shed light on the relative differences discussed above.

E. OUTLINE OF THIS PAPER

This paper is divided into three chapters. Chapter 1 provides a brief description of Urban HPAC and an overview of the JU03 field experiment. Chapter 2 provides details of our preparation of Urban HPAC predictions and descriptions of the methodologies used to compare predictions to observations. Chapter 3 presents the most important results of these comparative analyses. In addition to comparing varying urban dispersion modes of HPAC, the effect of differing meteorological inputs on predictions is discussed and comparisons with the results of previous Urban 2000 and MUST studies are highlighted. Finally, Appendix A presents a list of acronyms, Appendix B contains supplementary figures that depict comparative MOE results, and Appendix C provides an extract from the task order that supported this research.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

The potential effects of the atmospheric release of hazardous materials continue to be of concern to the nation – a concern especially acute in more densely populated urban areas. In addition, military activities increasingly are conducted in urban, highly populated areas. Therefore, the possible release of hazardous materials in an urban environment is especially troubling to both civilian and military authorities.

In the case of hazardous material releases, effective mitigation in urban settings will require an understanding of the transport and dispersion of these hazards in the urban environment. The DoD and DHS need to be able to estimate the effects of hazardous releases within an urban environment on the underlying population to aid planning, emergency response, and recovery efforts. These estimates require accurate knowledge of the concentrations of dispersed material in time and space. Especially desired are estimates of where and when relatively low-level human effects thresholds are exceeded. Models that predict the transport and dispersion of hazardous materials in urban settings are needed to develop doctrine, plan counter-proliferation operations, determine and characterize the source of the hazardous release based on limited observations, develop casualty estimates and emergency response plans, conduct forensic analyses, and perhaps, help guide near real-time emergency response activities.

Improved characterization and understanding of urban T&D will allow for more robust modeling. The buildings associated with urban regions, sometimes large and often closely packed, create their own roughness-induced boundary layers that can disperse toxic materials in ways that are not completely understood. Other features of the urban environment, including traffic-induced turbulence, heat island formation, flows associated with the deep canyons of some large cities, the relative lack of moisture, and differential heating on building faces, also can have varied effects on the transport and dispersion of hazardous materials.

Urban T&D modeling has been a subject of study since the late 1970s [Ref. 1-1]. While there are a relatively large set of atmospheric T&D models available for application to open field (rural area) hazardous gas releases, only a few models have been developed to simulate T&D in an urban environment.

Until recently, urban transport and dispersion models could be divided into two main categories: (1) low-fidelity models (e.g., urban canopy models) that account for the large-scale effects of urban terrain, such as drag from buildings and boundary layer perturbations, and (2) high-fidelity models (e.g., computational fluid dynamics models) that include detailed representations of buildings, streets, and other urban features, as well as the effects of traffic-induced turbulence, heat island formation, flows associated with the deep street canyons of some large cities, the relative lack of moisture, and differential heating on building faces. "In between" lies a recent class of urban T&D models that take into account detailed urban features, but employ empirical turbulence and wind profile parameterizations derived from these urban details. Examples are the particle-based MESO/RUSTIC model [Ref. 1-2] and the puff-based HPAC model [Ref. 1-3]. In this paper, we concentrate on the urbanized version of the HPAC model.

This paper describes comparisons of the tracer gas observations of the Joint Urban 2003 field experiment to HPAC predictions. This introductory chapter provides a brief description of the urbanized version of HPAC (Urban HPAC), a summary of previous evaluations of Urban HPAC, and a brief description of the Joint Urban 2003 field experiment (known henceforth as JU03). This introductory chapter also provides graphical examples of our analyses and processing of the observed concentration data.

A. URBAN HPAC DESCRIPTION

DTRA's HPAC is composed of a suite of software modules that can generate source terms for hazardous releases, retrieve and prepare meteorological information for use in a prediction, model the transport and dispersion of the hazardous release over time, and plot and report the results of these calculations [Ref. 1-3]. HPAC has been applied to various national defense problems, including military studies and operational planning.

For hazardous material transport and dispersion, HPAC uses the SCIPUFF model and an associated mean wind field model [Ref. 1-4]. SCIPUFF is a Lagrangian model for atmospheric dispersion that uses the Gaussian puff numerical method – an arbitrary time-dependent concentration field is represented by a superposition of three-dimensional Gaussian distributions – and bases its turbulent diffusion parameterization on second-order closure theories. This methodology provides a link between measurable velocity statistics and the predicted dispersion rates. This "second-order" feature implies that concentration fluctuation variance can also be computed, and this uncertainty estimate can be used as the basis for a probabilistic description of the dispersion prediction.

Given observations or predictions generated by a mesoscale meteorological model, HPAC can create mass-consistent wind fields that can be used to transport the hazardous material. Within HPAC two weather modules can be used to prepare these mass-consistent wind fields – SWIFT and MC-SCIPUFF. In the case of SWIFT, the underlying topography is required. For this study, the creation of HPAC predictions was completed using SWIFT when possible.[1] It should be noted that neither SWIFT nor MC-SCIPUFF account explicitly for wind speed profiles below the mean building height (i.e., within the urban canopy).

For urban applications of HPAC, the vertical wind and turbulence profiles can be modified to account for urban effects. The basic exponential profile shape used is described in Ref. 1-6 and has been shown to fit a variety of canopy types, including plant canopies as well as discrete objects. The canopy model has also been compared with experimental dispersion data within a plant canopy [Ref. 1-7] and with cube arrays in a wind tunnel [Ref. 1-8]. In using HPAC to provide predictions in an urban environment, one can conveniently capture some of the effects of the urban canopy on transport and dispersion by setting the surface type to "urban." This has the effect of setting the canopy height to 30 meters, surface roughness to 0.5 meters, the albedo to 0.18, and the Bowen ratio to 2.0. For the baseline Urban HPAC predictions examined in this study, the surface type was set to "urban." The terrain elevation associated with the Oklahoma City area – the location of JU03 – was also included in these baseline Urban HPAC predictions.

In addition to the baseline Urban HPAC predictive capability described above, also referred to as urban canopy (UC) mode, Urban HPAC offers an urban dispersion model (UDM) and an urban windfield module (UWM), either or both of which can be invoked. In order to use UDM and UWM, Urban HPAC requires a building database that provides the locations, planar geometries, and heights of buildings to support the calculation of flows in the urban regime. The newest feature of Urban HPAC (still being tested) is the Micro SWIFT SPRAY (MSS) module, which includes a particle-based Monte Carlo component.[2] Each of these urban T&D specific components is discussed below.

---

1 SWIFT is derived from the MINERVE - methode d'interpolation et de reconstitution tridimensionelle d'un champ de vent [Ref. 1-5].

2 SPRAY = a Monte-Carlo Lagrangian dispersion code.

1. Brief UDM Description

The UDM component of Urban HPAC was created by the United Kingdom's Defense Science and Technology Laboratory [Ref. 1-9] The UDM is designed to compute the transport and dispersion of an instantaneous discharge ("a puff" or train(s) of puffs) of pollutant released over a surface containing a mixture of open and urban areas. UDM considers variations in dispersion rates as a function of surface changes and direct interaction of the pollutant cloud with surface obstacles. UDM is based on ensemble mean Gaussian puff dispersion methodology but allows surface obstacles to modify the dispersion patterns. UDM empirical parameters are set based on extensive wind tunnel experiments.

Three regimes for calculation are identified within UDM: open regime, urban regime, and longer-range regime. Open regime conditions are satisfied for an obstacle density less than 5 percent (per unit area) and accounts for a dispersing puff interacting with individual obstacles, perhaps in succession. The urban regime is defined for obstacle densities greater than 5 percent and puff sizes on the order of the obstacle size. In this regime, the puff interacts with the relatively closely packed obstacles. The longer-range regime is assigned for obstacle densities greater than 5 percent and puff sizes large with respect to the obstacles. When no surface obstacles are present, as can occur in the open and longer-range regimes, HPAC (SCIPUFF) is used for the calculations within the integrated UDM–HPAC (i.e., Urban HPAC).

If a single obstacle is encountered in the open regime (of sufficient size as described previously), the puff will interact and be partitioned into entrained and unentrained fractions. The unentrained fraction will continue with some enhanced dispersion due to increased turbulence in the recovery region. The entrained fraction remains in a recirculating flow in the obstacle's wake, with a modeled characteristic residence time. As the entrained flow escapes, it is transported and dispersed separately from the unentrained fraction.

2. Brief UWM Description

The UWM predicts steady-state winds (speed and direction) inside the urban boundary layer using a canopy parameterization [Ref. 1-10]. UWM-generated average winds can then be used by Urban HPAC (for example, with or without UDM toggled on) to drive material transport and dispersion. UWM is meant to represent an improvement, for urban applications, over simply using SWIFT, which does not account for the effect of the buildings on the wind field.

UWM is a CFD code that is designed to provide a computationally fast solution, within a CFD framework, by considering spatially averaged obstacle effects. Therefore, the predicted winds of UWM represent spatio-temporal averages – spatial averages of the Reynolds-averaged equations of motion – which allows for fast solutions. This spatial averaging causes the momentum and energy equations to have extra terms that correspond to important physical processes. These processes include drag forces and turbulent energy production and are parameterized within UWM's canopy approach. This canopy parameterization approach assesses the drag and turbulent flux terms using a distributed drag model and an eddy diffusion model, respectively. The distributed drag model considers the area density of the obstacles and the mean height of the buildings, both provided within the HPAC software and based on the building geometry data of the included urban database. The drag coefficient is assumed to be constant. The turbulent flux terms are modeled following the procedures described for large-eddy simulation in Ref. 1-11.

Initial conditions for the UWM within HPAC can be set by including a mass-consistent, three-dimensional, gridded wind field based on observations (e.g., SWIFT) or by providing a numerical weather prediction.

3. Brief MSS Description

MSS is meant to provide a fast computation of the wind field within the urban environment while accounting for an exact representation of the buildings, e.g., as generated by Geographic Information Systems [Ref. 1-12]. As previously discussed, SWIFT can create a mass-consistent gridded wind field given topographic information and metrological inputs. Given information of the local buildings (locations, shapes, and sizes), Micro SWIFT creates a modified wind field by creating zones where the flow is modified according to the buildings' locations, and flow is adjusted to satisfy the continuity equation and impermeability on the ground and on the buildings' walls. The zones typically include a displacement zone, where the wind affects the building; a cavity zone, where flow may recirculate on the building side opposite of the wind direction; and a wake zone, where flow may be entrained for some time. Micro SWIFT also derives a diagnostic turbulence – diffusive coefficients and the turbulent kinetic energy dissipation rate – by considering the distance to the nearest obstacle as a mixing length and using this value for wind field local shear.

Micro SPRAY is a Lagrangian particle dispersion model, derived from SPRAY [Ref. 1-13], that can account for obstacles. Dispersion is simulated by following the

movement of a large number of fictitious particles, each representing a portion of the original released mass. The motions of the particles are obtained by applying an equation of motion that has two components – a mean component, which follows the local winds as defined by Micro SWIFT, and a stochastic component. The stochastic component of the particle motion follows a scheme described in Ref. 1-14 and includes a stochastic Gaussian term. Buildings are accounted for by Micro SPRAY by setting some of the cells of the terrain-following grid where the meteorological field is defined to be impermeable.

Within HPAC, MSS is used to compute hazardous material T&D over the first several hundred meters from the release point (at least as we employed MSS). After this distance, concentration information in the form of Gaussian puffs is "handed off" to SCIPUFF for the rest of the longer range calculation.

B. SUMMARY OF PAST EVALUATIONS OF URBAN HPAC PREDICTIONS

Two recent field experiments have been used to evaluate Urban HPAC predictions. First, a series of SF_6 gas releases were carried out in the Salt Lake City area in October 2000 (referred to as "Urban 2000") [Ref. 1-15]. This field trial had as its primary objective the collection of tracer concentrations and meteorological observations throughout an urban area in order to aid evaluations and further the development of atmospheric models. The Urban 2000 field trial examined releases of short duration (1 hour) from a point or short line (30 meters) source and in this sense simulated a notional terrorist attack in a U.S. city.

Meteorological and tracer measurements were conducted throughout the Salt Lake City urban region with an outermost arc of SF_6 samplers located 6 km downwind of the release. Six intensive operating periods (IOPs), which included SF_6 releases from within the downtown area, were associated with Urban 2000. For each of these IOPs, 3 separate 1-hour releases were monitored for 2 hours. Release times varied from 11 PM to 5 AM local time, that is, these SF_6 releases occurred at night or in the very early morning hours. During the time period of the Urban 2000 SF_6 releases, winds were generally from the east or southeast and varied from relatively light [averaging about 0.7 to 1.1 meters per second (m/s) at street level (1.5 meters)] – for four of the IOPs, to moderate (averaging 1.7 and 2.6 m/s at street level) for two of the IOPs [Ref. 1-16].

In 2001, a well-documented baseline (scaled down) urban setting was created in the desert of Utah and tracer gases were released. This atmospheric tracer and

meteorological study is known as the Mock Urban Setting Test (MUST) [Ref. 1-17]. The goal of MUST was to acquire meteorological and dispersion data sets at near full scale for use in urban dispersion model development and validation.

For the MUST experiment, a 12 by 10 array of shipping containers was positioned at Dugway Proving Ground (DPG), in Utah, to form an approximate 200-meter square, which was meant to represent a scaled version of an ideal, i.e., simple, urban environment. The (nominally identical) containers had dimensions 12.2 meters long, 2.42 meters wide, and 2.54 meters high. This 200-meter square area is smaller than a real urban setting but much larger than any wind tunnel experiment.

For MUST, the tracer gas propylene was released 68 times in September 2001. Five puff trials involved a series of instantaneous releases and 63 involved continuous release trials. Propylene concentrations were sampled at 50 Hz. Release locations within and just outside the array were varied; release heights were varied from 0.15 to 5.2 meters. The MUST releases were conducted at night or in the very early morning hours during stable atmospheric conditions. Additional information on the MUST experiment can be found in Ref. 1-17.

1. Urban HPAC – Urban 2000 Comparisons: Conclusions [Refs. 1-18 – 1-20]

In general, Urban HPAC over-predicted the observed concentrations and dosages[3] of Urban 2000. Average concentrations observed for the downtown area were over-predicted by about a factor of 1.67. Predictions of whether or not a relatively low threshold was exceeded (e.g., hazard regions) were substantially improved relative to predictions associated with absolute amounts of material – for all model modes that were examined. Urban HPAC model configurations that included the UDM configuration led to the best performing predictions. Addition of UWM to the UDM configuration did not lead to improvements relative to using the UDM configuration alone.

With respect to examinations of different meteorological input options, the best Urban 2000 predictions (of those examined) were generated by using either wind velocity information from an upwind profiler instrument (located about 5 km from the release location) or from using a mesoscale meteorological forecast. Two weather inputs that included meteorological information near the urban source resulted in relatively poor performance probably caused by including these more variable in-canopy observations as

---

[3] A dosage is defined here as a cumulative concentration over a defined time interval.

model input. Additional descriptions of conclusions associated with Urban HPAC evaluations using Urban 2000 can be found in Refs. 1-18 – 1-20.

2. Urban HPAC – MUST Comparisons: Conclusions [Refs. 1-21 and 1-22]

An important result of this previous study was the relative consistency of conclusions found between comparisons of Urban HPAC predictions of MUST observations and those just described for predictions of Urban 2000. For example, for both MUST and Urban 2000, improved predictions were associated with the inclusion of UDM; typically the HPAC baseline mode (urban canopy) led to larger false negative fractions with respect to hazard area predictions (i.e., number of samplers exceeding a threshold), and improvements associated with the inclusion of UWM were not found. We note that the MUST array is not simply a scaled down version of a typical city and, as such, corresponds to a relatively unique environment (e.g., relative to the Urban 2000 Salt Lake City experiment). One might view the MUST array scale as being most consistent with a small industrial facility. The above findings of relative consistency are therefore suggested to be somewhat robust with respect to urban environments.

Comparisons of Urban HPAC predictions and MUST observations also led to the conclusion that the use of meteorological observations well above, but directly over, the obstacle array – the 16 meter AGL sonic anemometer measurements – resulted in predictions with the best fit to the observed tracer concentrations.

Finally, the MUST field experiment, in large part because of its scale, served as a substantial challenge for HPAC and it has forced a close look at urban capabilities in ways that would not have been illuminated at actual urban scales. A few software integration errors associated with UDM and UWM were uncovered during this study that likely would not have been discovered at full scale. Additional descriptions of conclusions associated with Urban HPAC evaluations using MUST can be found in Refs. 1-21 – 1-22.

3. Some Future Questions Associated With Urban T&D Predictions

An integrated version of HPAC and MSS has only recently become available for independent evaluation. An important objective of the present study is to generate MSS-HPAC predictions of the Joint Urban 2003 field experiment and evaluate model performance via comparison with the observed concentrations. MSS was not available for the previous Urban 2000 and MUST studies.

It is also recognized that a wider range of meteorological conditions (perhaps higher wind speeds or more variable wind speeds and directions than were observed in Salt Lake City during Urban 2000 or during MUST) that should be available from the Joint Urban 2003 field experiment might create situations wherein the hoped-for benefits of UWM can be demonstrated. To this end, investigations of relatively high-resolution UWM predictions of the tracer gas releases conducted during Joint Urban 2003 may be of particular interest. In addition, the Joint Urban 2003 field experiment included releases during the day *and* at night, and therefore naturally led to a wider variety of atmospheric stability conditions than the previous Urban 2000 and MUST experiments, which were conducted at night (or during the very early morning hours). Therefore, exploring model performance variability as a function of atmospheric stability – "day versus night" is also recognized as an important component of this research.

Questions associated with what represents the most appropriate operational meteorological inputs for the prediction of transport and dispersion within an urban environment will, again, need to be addressed. An important component of the present Urban HPAC – Joint Urban 2003 comparisons, is to confirm, refute, or expand upon the previous findings. For example, Urban 2000 studies [Ref. 1-18] found that using close-in meteorological measurements, for example, from a downtown building top did not result in improved HPAC predictions. We speculate that some of the critical (that is, close to the source) weather measurements might have had significant spatio-temporal variability. For instance, for the building top measurements in Salt Lake City (Urban 2000) that were examined, weather direction measurements indicated relatively large temporal fluctuations – probably not too surprising for measurements done within a complex urban environment. The possible implication is that including such "within the urban environment meteorological observations" as model inputs might not necessarily improve model predictive performance relative to simply including upwind measurements. Perhaps better ways of applying meteorological observations taken within the complex urban environment are required.

In past studies [Refs. 1-18 and 1-21], it was speculated that including meteorological observations or vertical wind profiles above or upwind of an urban region might be a sufficient input to create reasonable HPAC hazard area predictions. Recently, it has been suggested that real-time wind field information from LiDAR (Light Detection and Ranging)-based systems can be used to construct vertical profiles near and above the city of interest to generate improved hazard predictions [Ref. 1-23]. These ideas can also

be explored using the observations of the Joint Urban 2003 field experiment and are the subject of continuing research.

It is expected that the results of this study will provide information that supports two broad Government decision-making areas: (1) identification of promising urban transport and dispersion codes to be further developed and potentially integrated into other military hazardous assessment systems (e.g., Joint Effects Model) and (2) specification of best practices with respect to using HPAC in urban environments to include urban mode settings and appropriate and sufficient use of meteorological inputs. In addition to the above, the analyses and results of this report support the general further development of hazardous material assessment tools.

C. BRIEF JOINT URBAN 2003 DESCRIPTION

Under the joint sponsorship of the DoD (DTRA) and DHS, a series of tracer gas releases were carried out in Oklahoma City starting on 28 June and ending on 31 July 2003. This field experiment, referred to as "*Joint Urban 2003*" [Ref. 1-23], represents a major urban study and included components that examined

- The urban atmospheric boundary layer [Refs. 1-24 and 1-25]
- Flows within a street canyon to include traffic induced turbulence [Refs. 1-26 through 1-30]
- Flows within, and downwind (to 4 km), of the tall-building core [Ref. 1-31]
- Surface energy balance within an urban area [Refs. 1-32 and 1-33]
- The dispersion of tracer into, out of, and within buildings.

We studied a reduced set of these components as described below.

1. Joint Urban 2003 (JU03) Sulfur Hexafluoride (SF_6) Releases

During 10 IOPs that covered the 30 days between 29 June 2003 and 28 July 2003, the tracer gas SF_6 was released in downtown Oklahoma City. In total, twenty-nine 30-minute continuous SF_6 releases were accomplished with 2 hours of sampler monitoring following the start of each release. An additional 40 instantaneous releases, also referred to as "puff" releases, were conducted during the 10 IOPs. In addition to monitoring the outdoor concentrations, infiltration of tracer gas into four downtown buildings was

studied during four of the IOPs. This initial study examines the 29 continuous SF_6 releases along with the observed outdoor concentrations.[4]

As shown in Figure 1-1, four release locations, all in downtown Oklahoma City, were used for the 29 continuous SF_6 releases, referred to here as Park Avenue (PA), Botanical Gardens (BG), Westin-A (WA), and Westin-B (WB). Table 1-1 provides summary information for each of the 29 continuous SF_6 releases that were examined in this study. The duration of each of these releases was 30 minutes. The altitude for each of the releases was about 1.9 meters AGL. Release rates varied from 1.9 to 5.0 grams per second (g/s), implying total released SF_6 masses of between 3.42 and 9.00 kilograms (kg) per release.

Figure 1-1. Overhead View of Downtown Oklahoma City with SF_6 Release Locations Shown

A garden hose with 0.625 inches (~ 1.6 cm) exit diameter was used for the continuous SF_6 releases. When using HPAC to simulate these releases, a two-dimensional (y- and z-direction) Gaussian spread can be used to define the detailed source term. We assumed that the Gaussian spread is equal to the radius of the hose. Therefore, we defined the Gaussian spread for the continuous source releases as 0.8 cm for all IOPs (both in y- and z-directions).

---

[4] Future analyses are proposed to consider the puff releases and comparative studies of indoor air concentrations are under consideration.

Table 1-1. Summary of JU03 Continuous SF_6 Releases [Extracted, in part, from Ref. 1-34.]

IOP # / Release #	2003 Date	Time of Day of Release (CDT)[5] – Day/Night	Release Location (WA, WB, BG, PA)	Release Rate (g/s)
1/1	29 June	1100 – Day	WA	4.8
1/2	29 June	1300 – Day	WA	4.9
2/1	2 July	1100 – Day	WB	5.0
2/2	2 July	1300 – Day	WB	5.0
2/3	2 July	1500 – Day	WB	5.0
3/1	7 July	1100 – Day	BG	5.0
3/2	7 July	1300 – Day	BG	3.0
3/3	7 July	1500 – Day	BG	3.0
4/1	9 July	1100 – Day	BG	3.1
4/2	9 July	1300 – Day	BG	3.0
4/3	9 July	1500 – Day	BG	3.0
5/1	13 July	0900 – Day	BG	2.2
5/2	13 July	1100 – Day	BG	3.0
5/3	13 July	1300 – Day	BG	3.1
6/1	16 July	900 – Day	BG	3.0
6/2	16 July	1100 – Day	BG	3.2
6/3	16 July	1300 – Day	BG	3.0
7/1	19 July	2300 – Night	BG	3.0
7/2	19 July	0100 – Night	BG	2.0
7/3	19 July	0300 – Night	BG	2.0
8/1	25 July	2300 – Night	WB	3.1
8/2	25 July	0100 – Night	WB	3.0
8/3	25 July	0300 – Night	WB	3.0
9/1	27 July	2300 – Night	PA	2.0
9/2	27 July	0100 – Night	PA	2.0
9/3	27 July	0300 – Night	PA	2.1
10/1	29 July	2300 – Night	PA	2.2
10/2	29 July	0100 – Night	PA	1.9
10/3	29 July	0300 – Night	PA	2.2

[5] CDT = Central Daylight Time.

Inspection of Table 1-1 shows that of the 29 releases, 17 occurred during the day and 12 occurred at night. For the four release locations – WA, WB, BG, and PA – there were two (day), six (3 day, 3 night), fifteen (12 day, 3 night), and six (night) releases, respectively.

Extensive meteorological measurements were made during the JU03 experiment. The meteorological measurements that were made and that are most important to this study are discussed in Chapter 2.

2. Samplers Used for This Study

This study focused on comparisons of predicted and observed concentrations near the surface. For the majority of this study, we used the stationary SF_6 sampling observations associated with the Programmable Integrating Gas Samplers (PIGS) and enhanced PIGS, referred to as Super PIGS [Ref. 1-35].[6] The PIGS and Super PIGS were operated by personnel from the National Oceanographic and Atmospheric Administration's (NOAA) Air Resources Laboratory (ARL) Field Research Division (FRD). A single PIGS includes 12 microprocessor-controlled air pumps capable of filling 12 bags using the same programmed fill times. For example, if fill times are set at 10 minutes, a single PIGS could cover 2 hours at a time resolution of 10 minutes for the entire 2-hour period. The Super PIGS includes 12 microprocessor-controlled valves and one pump and can fill 12 bags using different programmed times. For example, for close-in samplers, one might use a finer time resolution just after the release and a courser resolution later. Sampler bag integration time varied from 5 to 30 minutes. An automated tracer gas analysis system that includes a gas chromatograph and an electron capture detector was used to assess concentrations in ranges between 2 parts per trillion by volume (pptv or simply ppt) and about 200,000 pptv [Ref. 1-35].[7]

Figure 1-2 shows the typical locations of the SF_6 ground-based samplers ("ARL bag samplers")[8] that were used for this study. These "ground-based" samplers were located at approximately 3 meters AGL, typically mounted on light poles. Fifty-five ARL bag samplers were located in the central business district (CBD, i.e., the city center

---

6 Additional discussion of concentration measurements, including other samplers that were used, can be found in Ref. 1-23.

7 Detailed sample collection and tracer gas analysis techniques, including extensive quality control procedures, are described in Ref. 1-35.

8 We refer to the combination of PIGS and Super PIGS here as the ARL bag samplers.

near the release points) and 23 were located on an arc located about 1 km from the release point. Similarly, 21 ARL bag samplers were located on a 2 km arc, and 21 were located on the 4 km arc. In addition, 10 bag samplers were located on rooftops and 4 were located within tunnels. These rooftop and tunnel-based samplers were not examined for this initial comparative study. The sampler locations shown in Figure 1-2 would best accommodate a southeast wind direction (i.e., winds coming from the southeast). During the field experiment, sampler locations were sometimes moved between IOPs to accommodate the prevailing wind direction, e.g., the orientation of the arc samplers shown in Figure 1-2 was sometimes rotated by about 45 degrees to best accommodate winds out of the south.

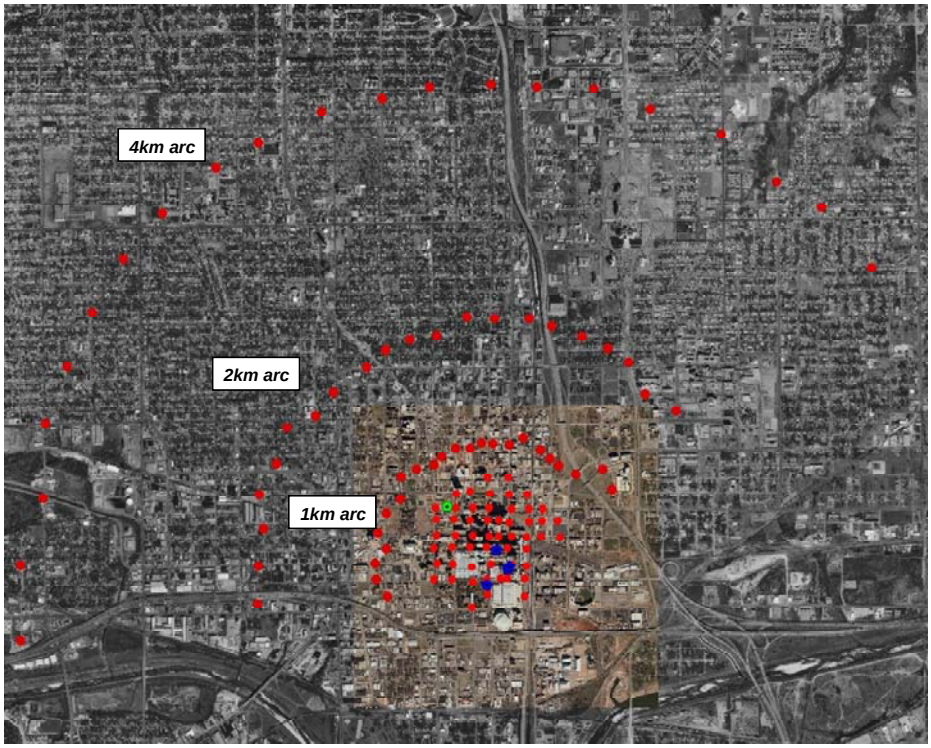


Figure 1-2. Locations of ARL Bag Surface Samplers

The blue pentagons correspond to the locations of the IOP releases (Figure 1-1) and the green pentagon corresponds to the location of a single mini-IOP that was designed to examine vertical dispersion at a single location – on the "Crane-mounted" samplers – discussed later in this section.

The background level of SF_6 for JU03 was 5 pptv as determined by ARL FRD [Ref. 1-36]. The following "background" protocol was followed for this analysis. First, 5 pptv was added to all predictions before conducting comparisons with observations. Next, any observation that was observed as below 5 pptv was set to exactly 5 pptv. Some values in the distributed set of concentration observations were denoted as unusable for

any of several quality-control-related reasons. These observations, and their paired predictions, were removed from this comparative analysis. Finally, pptv measurements were transformed to values with kg/m^3 units using a conversion factor of 5.803 $\times 10^{-12}$ kg/m^3 = 1 pptv.[9]

A variety of graphical representations of predictions and observations were created for all 29 releases and used throughout this study. Examples of some of the graphical representations of the observed concentrations that were created for all releases are shown in Figures 1-3 through 1-8. Figures 1-3 and 1-4 present average 30-minute observed concentrations during IOP 9, Release 1 in the CBD for the four 30-minute time periods that were monitored. Figure 1-5 presents the average 2-hour concentration for this release. In order to generate a 2-hour average concentration for a particular sampler location, we required that samplers be monitored and useable for all 30-minute time periods, that is, no interpolation in time was used to create the 2-hour average concentrations. For example, some of the CBD samplers were not used after the first hour and so 2-hour average concentrations could not be computed at these locations. This is evident in the "holes" in the CBD shown in Figures 1-4 and 1-5.

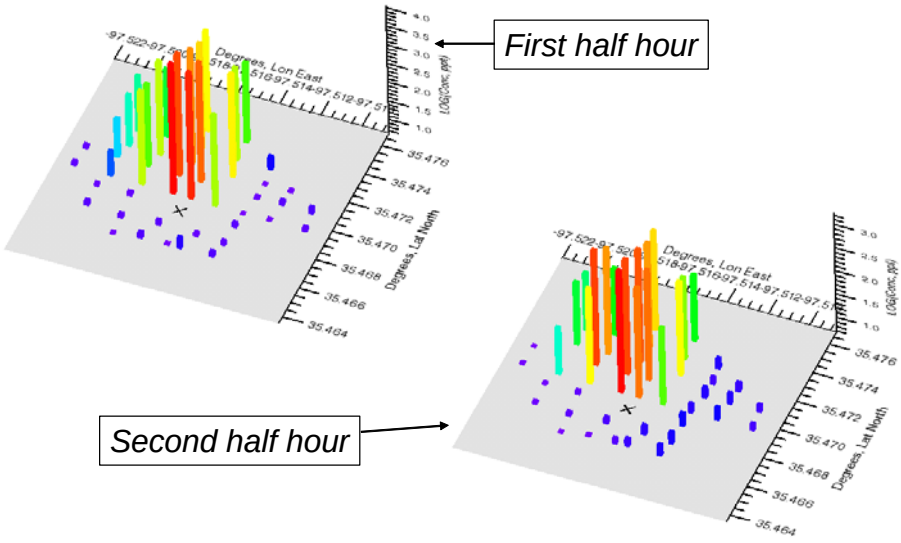


Figure 1-3. Surface Sampler Concentration Observations for the 1st and 2nd Half Hour After IOP 9, Release 1: 30-minute Average Concentrations in the Central Business District (CBD)

Z-axis is in log(concentration pptv) and the colors purple, blue, green, yellow, orange, and red also encode the magnitude of the concentration; the "x" marks the release location.

---

9 This conversion factor was chosen to be consistent with the exact factor used by ITT so that future comparisons with ITT- or IDA-created MESO/RUSTIC predictions would be unbiased.

Figures 1-6 through 1-8 show similar plots, again for IOP 9, Release 1, for the samplers located on the ~1, ~2, and ~4 km arcs. In this case, sampling was accomplished on each arc for each 30-minute time period. Therefore, the 2-hour average concentration observations are relatively complete for the arc-based samplers.

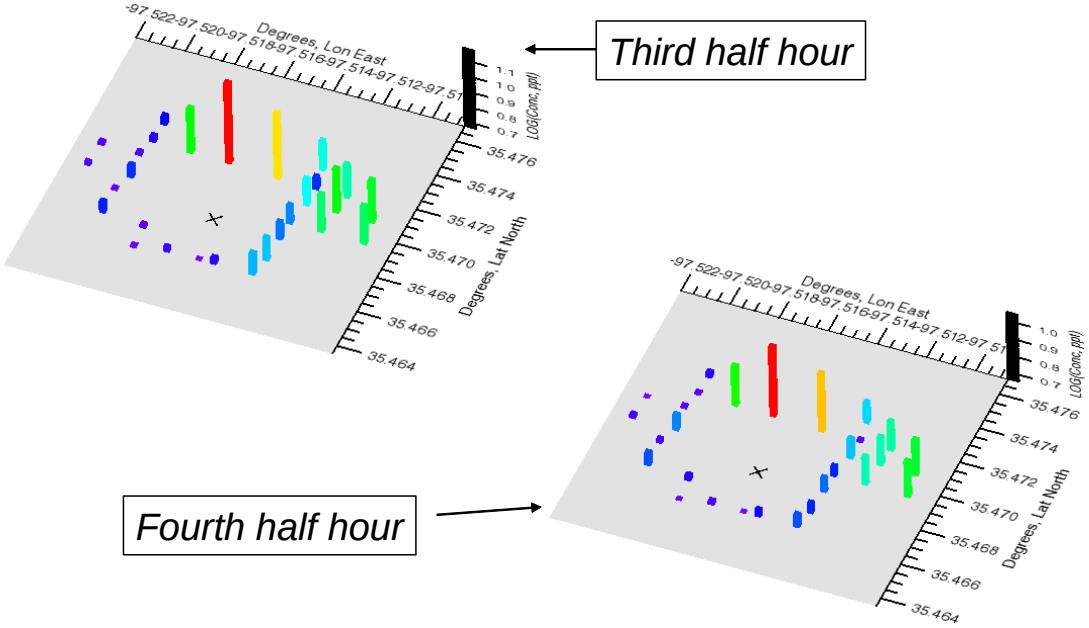


Figure 1-4. Surface Sampler Concentration Observations for the 3rd and 4th Half Hour After IOP 9, Release 1: 30-minute Average Concentrations in the Central Business District (CBD)

Z-axis is in log(concentration pptv) and the colors purple, blue, green, yellow, orange, and red also encode the magnitude of the concentration; the "x" marks the release location.

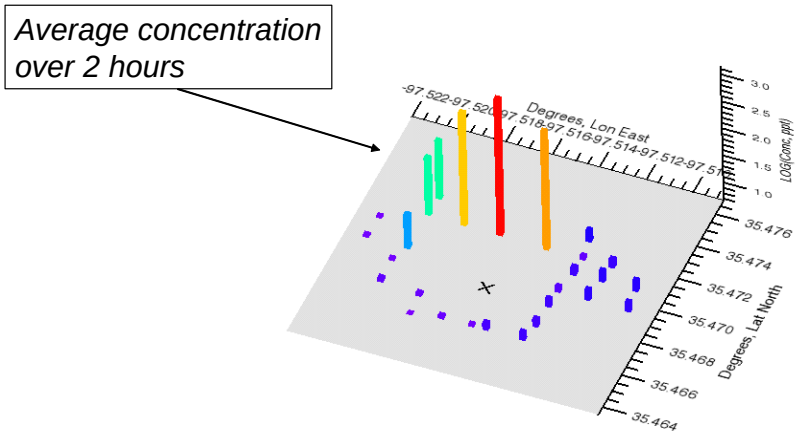


Figure 1-5. Surface Sampler Concentration Observations for IOP 9, Release 1: 2-hour Average Concentrations in the Central Business District (CBD)

Z-axis is in log(concentration pptv) and the colors purple, blue, green, yellow, orange, and red also encode the magnitude of the concentration; the "x" marks the release location.

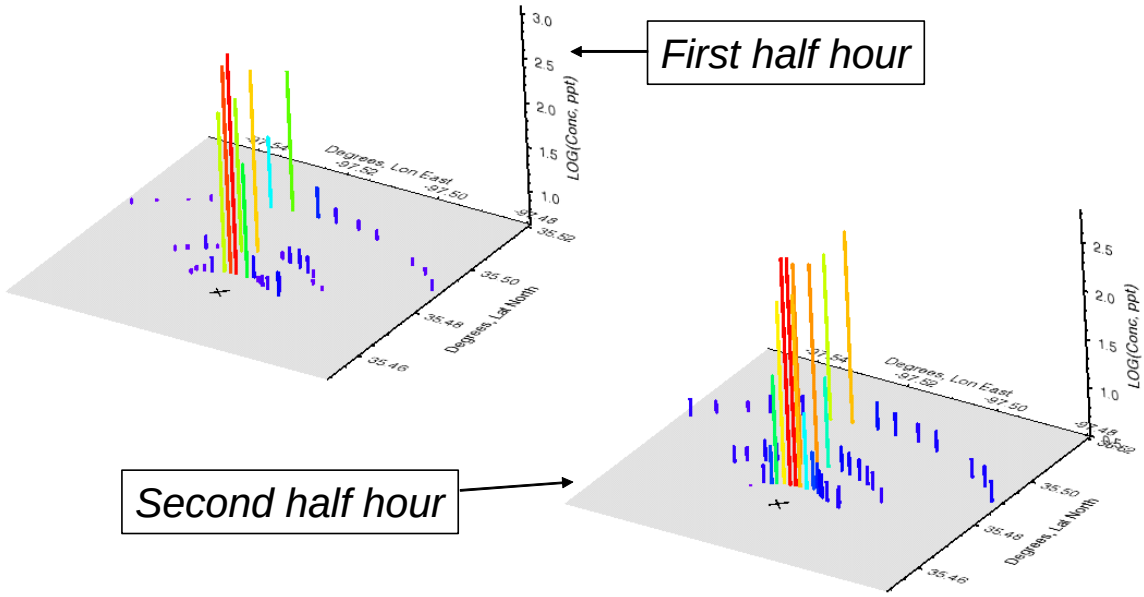


Figure 1-6. Surface Sampler Concentration Observations for the 1st and 2nd Half Hour After IOP 9, Release 1: 30-minute Average Concentrations on the Arcs

Z-axis is in log(concentration pptv) and the colors purple, blue, green, yellow, orange, and red also encode the magnitude of the concentration; the "x" marks the release location.

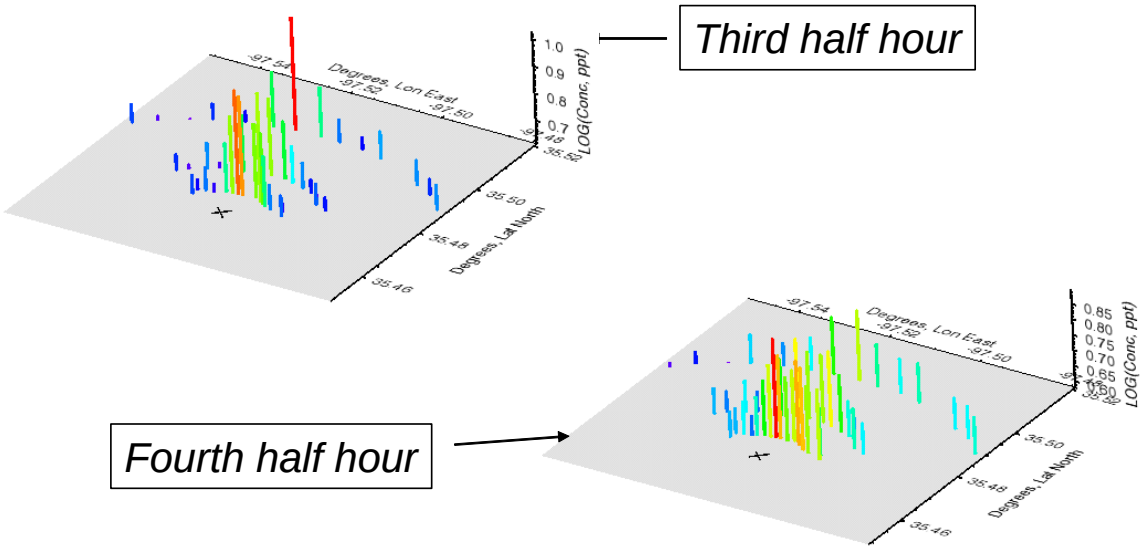


Figure 1-7. Surface Sampler Concentration Observations for the 3rd and 4th Half Hour After IOP 9, Release 1: 30-minute Average Concentrations on the Arcs

Z-axis is in log(concentration pptv) and the colors purple, blue, green, yellow, orange, and red also encode the magnitude of the concentration; the "x" marks the release location.

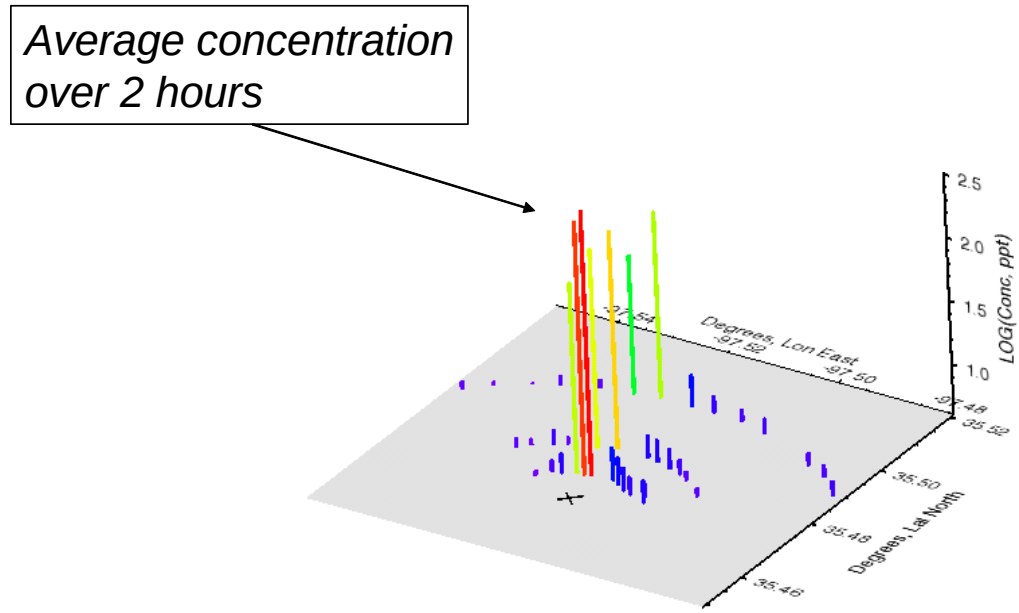


Figure 1-8. Surface Sampler Concentration Observations for IOP 9, Release 1: 2-hour Average Concentrations on the Arcs

Z-axis is in log(concentration pptv) and the colors purple, blue, green, yellow, orange, and red also encode the magnitude of the concentration; the "x" marks the release location.

Given values at some discrete (perhaps irregular) set of samplers, the process of (area) interpolation provides intermediate values on some regular grid of points. The resulting regular grid of functional values could be used to obtain contours of "hazard" areas (areas within a critical threshold contour). Interpolation procedures can be carried out either in linear or logarithmic space. When interpolating actual plume concentrations or dosages varying over orders of magnitude, one might favor interpolation schemes in logarithmic space.

The Delaunay triangulation procedure is useful for the interpolation, analysis, and visual display of irregularly, discretely gridded data. From a set of discrete points (sampler coordinates), a planar triangulation is formed, satisfying the property that the circumscribed circle of any triangle in the triangulation contains no other vertices in its interior.[10] For any point that is within some triangle (formed via Delaunay triangulation), a linear interpolation routine using values at the vertices of the triangle is used to calculate the value at that point. Delaunay triangulation is efficiently implemented in Interactive Data Language (IDL) software and forms a core interpolation routine for

[10] The Delaunay triangulation procedure is closely related to the procedure followed to create Voronoi polygons [Ref. 1-37].

display of irregularly gridded data. Delaunay triangulation has been previously used as the basis for area interpolation of observed concentration data [Ref. 1-38].

We used the above procedure in the following way. First, we transformed the data (observations and predictions) logarithmically and then followed the above procedure. The displays reported in Figure 1-9 are based on the logarithmic transformation of the data followed by Delaunay triangulation and linear interpolation as described above.[11] The adopted procedure, while simple and yielding some perhaps less visually pleasing sharp edges, appeared to be robust and necessarily maintains the actual observed values at the sampler locations, which would not be true for many fitting procedures.[12]

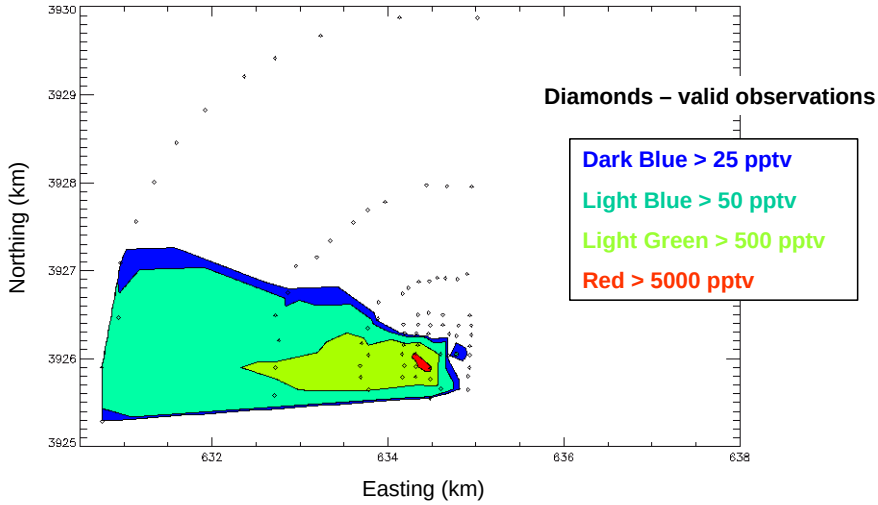


Figure 1-9. 1-Hour Average Concentration Contours for the 1st Hour of IOP 1, Release 2 Based on Delaunay Interpolation

Figure 1-10 shows the concentration contours that result from interpolation for the four 3-minute time periods that were monitored and provide results for all surface samplers and for just the CBD ("zoomed obs"). Figures 1-11 and 1-12 show similar plots and include comparisons with predictions. In the case of the predictions, the

---

11 All plots were done with a resolution of 501 × 501 points (i.e., 500 by 500 delta x, delta y intervals). For the complete plots (e.g., Figure 1-9), xrange = 7.5 km, dx = 0.015 km, yrange = 5.0 km, and dy = 0.010 km. For the zoomed plots (e.g., Figure 1-10 "Zoomed"), xrange = 1.9 km, dx = 0.0038 km, yrange = 1.5 km, and dy = 0.003 km. For all of the contoured plots that follow, UTM coordinates are used with the horizontal axis in km east ("Easting") and the vertical axis in km north ("Northing"). These headings were left off of most of the plots that follow to make them less cluttered.

12 A few other interpolation techniques were briefly examined in Ref. 1-39.

interpolation and formation of contours was done exactly the same was as it was for the observations.

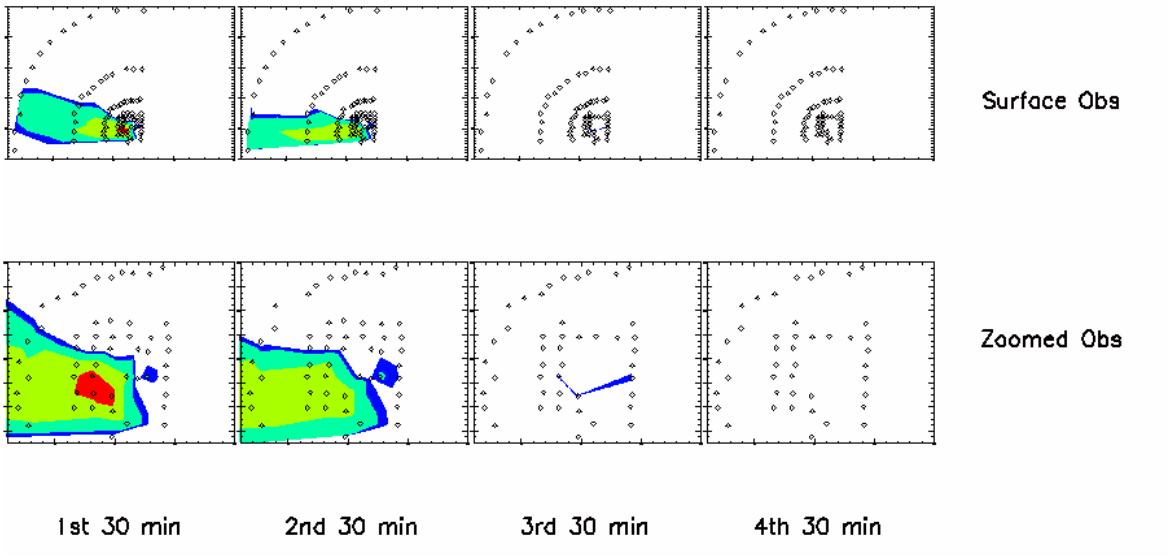


Figure 1-10. 30-Minute Average Concentration Contours for the Four 30-Minute Time Periods That Were Monitored During IOP 1, Release 2

The thin blue triangles shown in the 3rd 30-minute segment reflect the actual results of interpolation, i.e., no smoothing of the results were done for these plots.

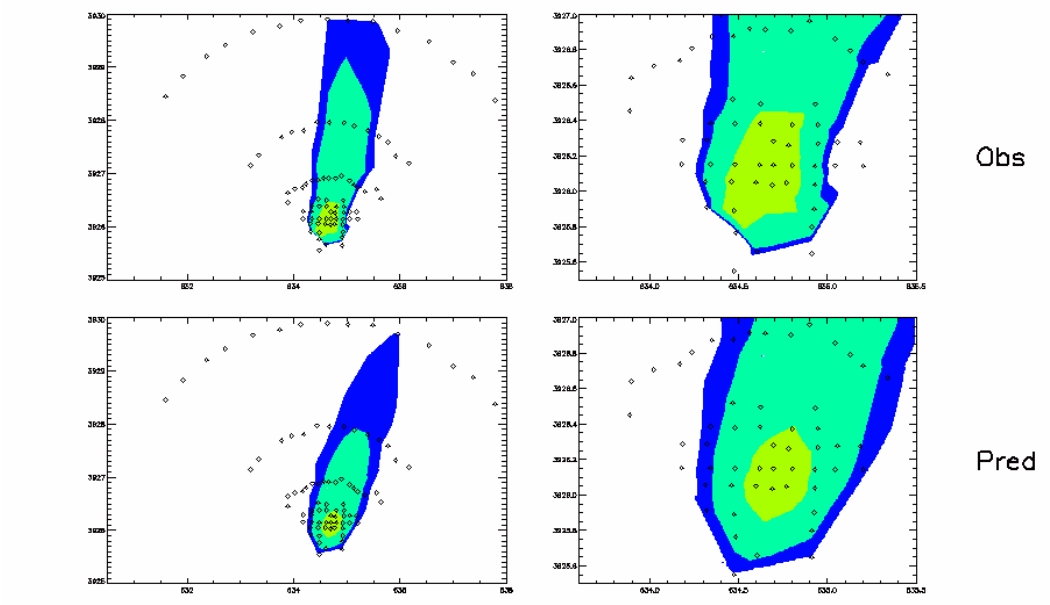


Figure 1-11. 1-Hour Average Concentration Contours for the 1st Hour of IOP 6, Release 2: Surface Observations and Predictions at the Surface Samplers

The predictions are based on the urban canopy (or baseline) mode of HPAC and use the baseline meteorological input option (to be described in Chapter 2).

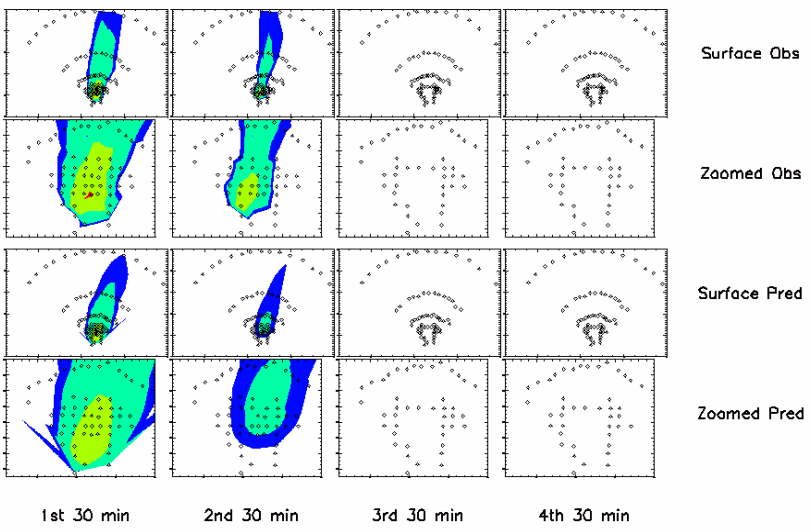


Figure 1-12. 30-Minute Average Concentration Contours for the Four 30-Minute Time Periods That Were Monitored During IOP 6, Release 2: Surface Observations and Predictions at the Surface Samplers are Shown for All Surface Samplers and for the CBD-"Zoomed"

The predictions are based on the urban canopy (or baseline) mode of HPAC and use the baseline meteorological input option (to be described in Chapter 2).

In addition to the PIGS/Super PIGS (ARL bag samplers) described previously, a few other sets of samplers were available during JU03. Some samplers were mobile (Lawrence Livermore National Laboratory "blue boxes" and Argonne Research Laboratory vans), some provided continuous (that is, very high frequency) measurements [Miniature Infrared Analyzer (MIRAN) and ITT continuous samplers], and sometimes additional samplers were placed close to the release or at various locations within the street canyon, on rooftops, or on a crane (green pentagon in Figure 1-2) to provide concentration measurements at a variety of altitudes at the same location. Typically, these samplers were not always available and/or only available in certain locations. The concentrations measured by the MIRAN and ITT continuous samplers were briefly examined. In particular, "movies" (animations using the audio video interleave –".avi"– file format) were created that show the concentration time history of the MIRAN and ITT measurements.[13] Figure 1-13 provides example "snapshots" from these movies. These movies were made using 10-second time-reversed concentration averages, and the areas of the circles plotted in Figure 1-13 are proportional to the concentration averages.

---

[13] These ".avi" files are available upon request. The associated MIRAN and ITT high-frequency concentration observations were not used in this comparative analysis because they were only available in limited locations in the CBD. Rather, and in part to ensure sampling protocol consistency, we relied on the ARL bag surface samplers for our comparative analyses.

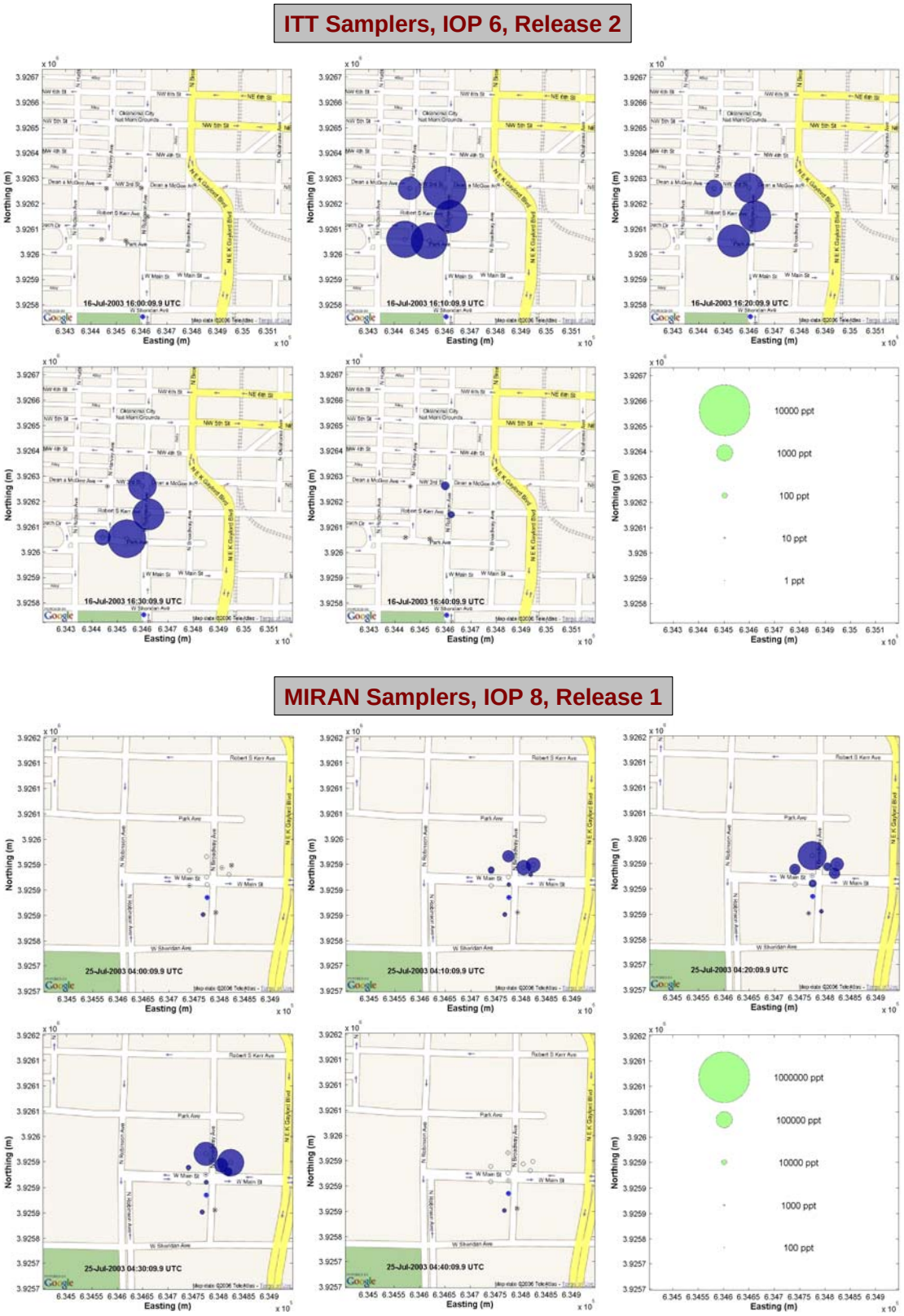


Figure 1-13. Example Concentration Observation Snapshots from the ITT (IOP 6, Release 2) and MIRAN (IOP 8, Release 1) Continuous Samplers

The ARL (bag) surface samplers associated with 29 releases from 10 IOPs (see Table 1-1) were the source of concentration observations used for comparative analysis in this study. In addition, we examined the concentration measurements made on the crane at a variety of altitudes at the same location. On 15 July 2003, a mini-IOP was conducted that had as an important goal the placement of significant tracer gas concentrations on the samplers located on the crane (at various altitudes). In general, we found only weak concentration dependence on altitude for the relatively low altitudes that could be investigated with the crane. Figure 1-14 shows the observed concentration time history for the crane-mounted samplers as a function of sampler height. Figure 1-15 shows observations for the first 30 minutes after the first release of the mini-IOP at each of the monitored levels. We did not use the observations from these crane-mounted samplers for comparisons in this study – that is, our comparisons of model predictions were limited to surface samplers only and, as such, provide a relative assessment of the models' abilities to predict concentrations in the operationally important near-ground region.

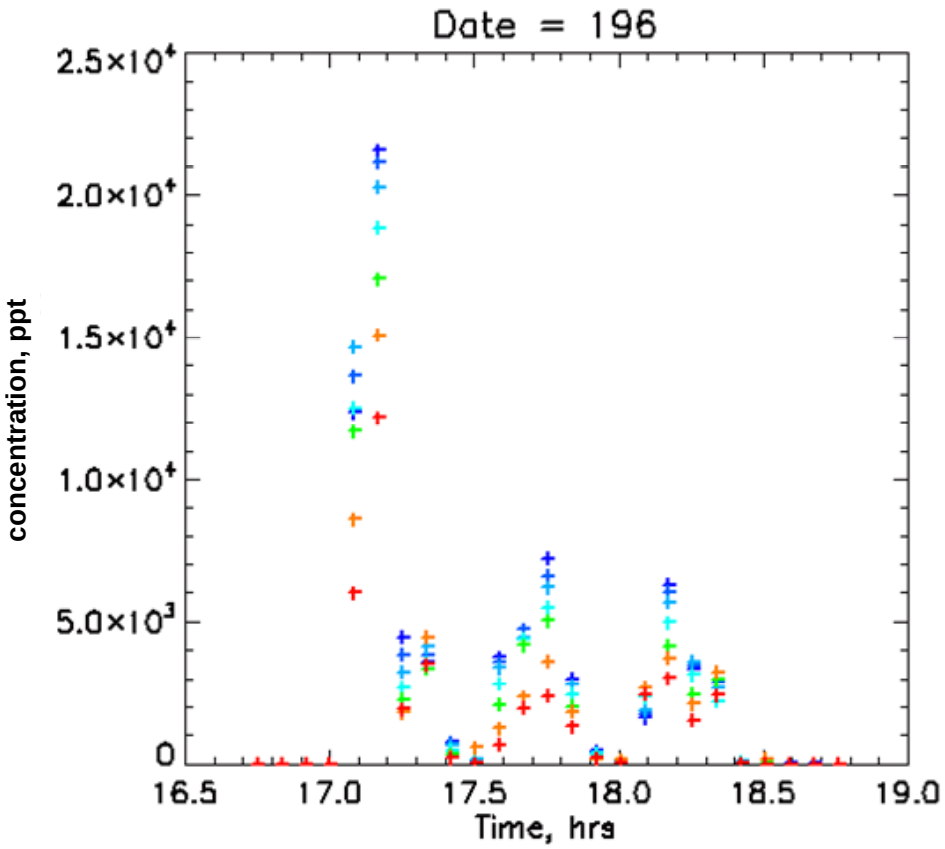


Figure 1-14. Concentration Observations from Crane-Mounted Samplers During the Monitoring Period for the Three Releases of the 15 July 2003 Mini-IOP

Altitude levels in meters: dark blue = 10.7, mid-level blue1 = 17.5, mid-level blue 2 = 24.4, light blue = 34.7, green = 48.4, orange = 62.1, and red = 75.8.

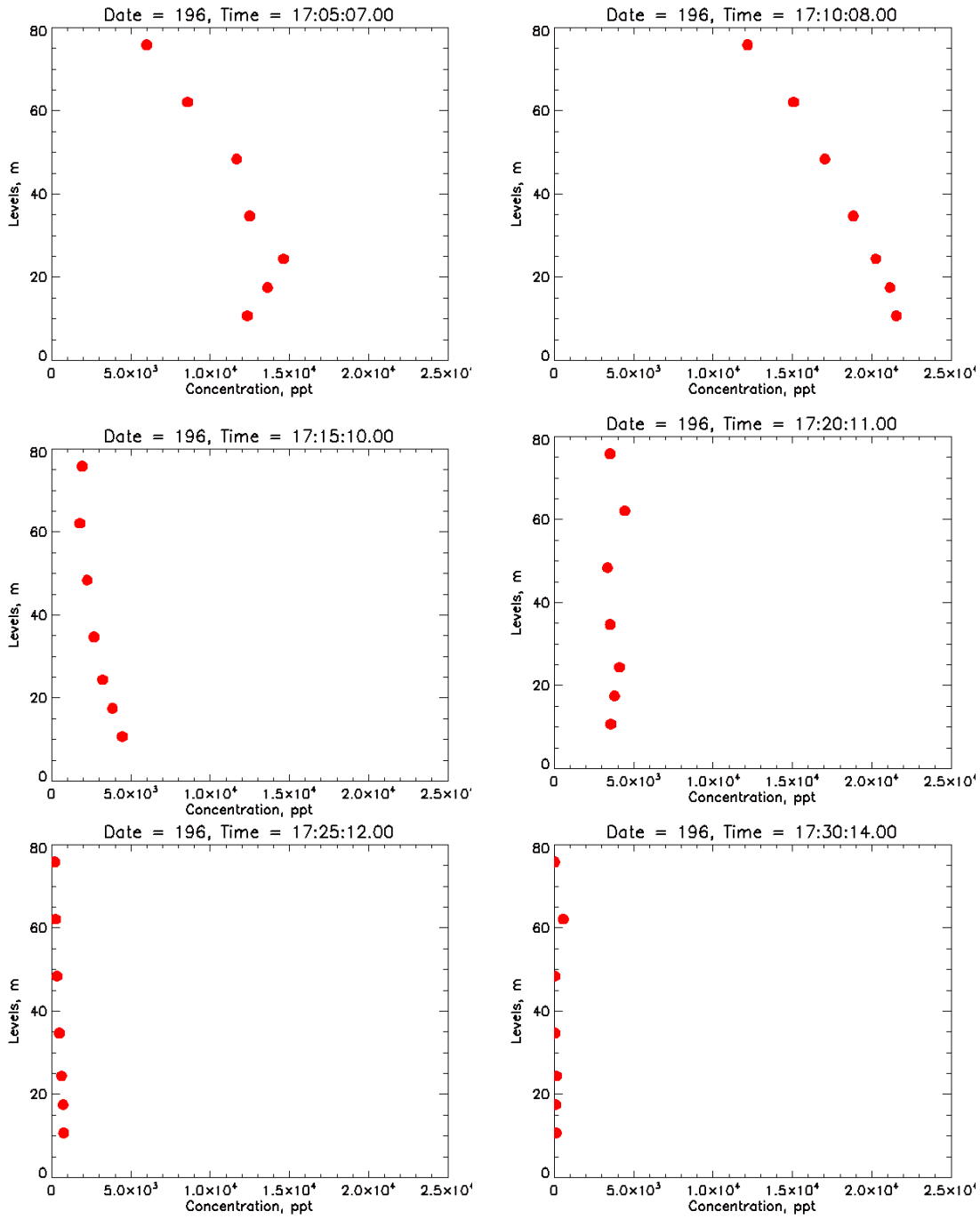


Figure 1-15. Concentration Observations from Crane-Mounted Samplers During the Monitoring Period of the First Release of the 15 July 2003 Mini-IOP

The "levels" on the y-axis correspond to the height AGL of the sample collection.

The next chapter describes the preparation of JU03 predictions and the methodologies used to compare predictions and observations.

REFERENCES

1-1. Britter, R. E. and S. R. Hanna, 2003: Flow and dispersion in urban areas. *Ann. Rev. of Fluid Mech.,* **35**, 469-496.

1-2. a) Diehl, S. R., D. T. Smith, and M. Sydor, 1982: Random-walk simulation of gradient-transfer processes applied to dispersion of stack emission from coal-fired power plants. *J. Appl. Meteor.,* **21**, 69-83., b) Hendricks, E., D. A. Burrows, S. Diehl, and R. Keith, 2004: Dispersion in the downtown Oklahoma City domain: comparisons between the Joint Urban 2003 data and the RUSTIC/MESO models. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27, and c) Burrows, D. A., R. Keith, S. Diehl, and E. Hendricks, 2004: A fast running urban air flow model. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-3. DTRA, 2001: The Hazard Prediction and Assessment Capability (HPAC) user's guide version 4.0.3. Prepared for the Defense Threat Reduction Agency by Science Applications International Corporation, Rep. HPAC-UGUIDE-02-U-RAC0, 602 pp.

1-4. Sykes, R. I., S. F. Parker, and R. S. Gabruk, 1996: SCIPUFF – a generalized hazard prediction model. *Preprints, Ninth Joint Conf. on the Applications of Air Pollution Meteorology, Atlanta, GA, Amer. Meteor. Soc.,* 184-188.

1-5. ARIA Technologies, 2001: General design manual, MINERVE Wind Field Model, version 7.0. ARIA Technologies, 72 pp.

1-6. Cionco, R. M., 1972: A wind-profile index for canopy flow. *Boundary-Layer Met.,* **3**, 255-263.

1-7. Aylor, D. E. and F. J. Ferrandino, 1989: Dispersion of spores released from an elevated line source within a wheat canopy. *Boundary-Layer Met.,* **46**, 251-273.

1-8. Macdonald, R. W., 2000: Modelling the mean velocity in the urban canopy layer. *Boundary-Layer Met.,* **97**, 25-45.

1-9. Hall, D. J., A. M. Spanton, I. H. Griffiths, M. Hargrave, and S. Walker, 2002: The Urban Dispersion Model (UDM): Version 2.2 Technical Documentation. DSTL/TR04774, 106 pp. [Available from Defense Science and Technology Laboratory, Porton Down, Salisbury, SP4 0JQ, UK.]

1-10. Lim, D., D. S. Henn, and R. I. Sykes, 2003: UWM Version 0.1. Prepared for the Defense Threat Reduction Agency by Titan Research and Technology Division, Titan Corporation, Tech. Doc., 42 pp.

1-11. Sykes, R. I. and D. S. Henn 1989: Large-eddy simulation of turbulent sheared convection. *J. Atmos. Sci.,* **46**, 1106-1118.

1-12. Moussafir, J., O. Oldrini, G. Tinarelli, J. Sontowski, C. M. Dougherty, 2004: *Proceedings of the 9th Int'l Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes,* Garmisch-Partenkirchen, Germany, 1-4 June 2004, pp 114-118.

1-13. Tinarelli, G., D. Anfossi, G. Brusasca, E. Ferrero, U. Giostra, M. G. Morselli, J. Moussafir, F. Tampieri, and F. Trombetti, 1994: Lagrangian particle simulation of tracer dispersion in the lee of a schematic two-dimensional hill. *J. Appl. Meteor.,* **33**, 744-756.

1-14. Thomson, D. J., 1987: Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.,* **180**, 529-556.

1-15. Allwine, K. J., J. H. Shinn, G. E. Streit, K. L. Clawson, and M. Brown, 2002: Overview of URBAN 2000, A multiscale field study of dispersion through an urban environment. *Bull. Amer. Meteor. Soc.*, **83**, 521-536.

1-16. Hanna, S. R., R. Britter, and P. Franzese, 2003a: A baseline urban dispersion model evaluated with Salt Lake City and Los Angeles tracer data. *Atmos. Environ.,* **37**, 5069-5082.

1-17. Biltoft, C. A., 2002: *Customer report for Mock Urban Setting Test,* DPG Document No. WDTC_FR-01-121, Meteorology and Obscurants Division, West Desert Test Center, U.S. Army Dugway Proving Ground, Dugway Utah, 84022-5000.

1-18. Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of transport and dispersion model predictions of the URBAN 2000 field experiment. *J. Appl. Meteor*., **43**, 829-846.

1-19. a) Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of Urban 2000 observations and urban HPAC predictions paired in space and time. *Proceedings of the 9th Int'l Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, 1-4 June 2004, pp 172-176. b) Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of Urban 2000 observations and urban HPAC predictions paired in space and time. *72nd Military Operations Research Society Symposium*, Monterey, California, 22-24 June 2004. c) Warner, S. and N. Platt, 2003: *Analyses in Support of Initial Validation of Urban HPAC: Comparisons to Urban 2000 Observations,* IDA Document D-2870N. [Available from Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.]. d) Warner, S. and N. Platt, 2003: Comparisons of Urban 2000 Observations and urban HPAC predictions paired in space and time, part 1: methodology and measures. *Seventh Annual GMU Transport and Dispersion Modeling Workshop*, at George Mason University, 17-19 June 2003. e) Platt, N. and S. Warner, 2003: Comparisons of urban 2000 observations and urban HPAC predictions paired in space and time, part 2: results and conclusions. *Seventh Annual GMU Transport and Dispersion Modeling Workshop*, at George Mason University, 17-19 June

2003. f) Boybeyi, Z., J. C. Chang, S. R. Hanna, P. Franzese, N. Platt, and S. Warner, 2003: *Independent Evaluation of Urban HPAC with the Urban 2000 Field Data.* Prepared for the Defense Threat Reduction Agency.

1-20. Chang, J. C., S. R. Hanna, Z. Boybeyi, and P. Franzese, 2005: Use of Salt Lake City URBAN 2000 field data to evaluate urban Hazard Prediction Assessment Capability (HPAC) dispersion model. *J. Appl. Meteor*., **44**, 484-501.

1-21. Warner, S., N. Platt, J. F. Heagy, J. E. Jordan, and G. Bieberbach, 2006: Comparisons of transport and dispersion model predictions of the mock urban setting test field experiment. *J. Appl. Meteor. and Climatology*, **45**, 1414-1428.

1-22. Warner, S., N. Platt, J. F. Heagy, J. E. Jordan, and G. Bieberbach, 2005: Comparisons of transport and dispersion model predictions of the Mock Urban Setting Test (MUST) field experiment. Institute for Defense Analyses Paper P-4030. [Available from Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, VA 22311-1882.]

1-23. Allwine, K. J., M. J. Leach, L. W. Stockham, J. S. Shinn, R. P. Hosker, J. F. Bowers, and J. C. Pace, 2004: Overview of joint urban 2003—An atmospheric dispersion study in Oklahoma City. *Symp. on Planning, Nowcasting and Forecasting in the Urban Zone*, Seattle, WA, Amer. Meteor. Soc., January 12-16.

1-24. Berg, L. K., S. F. J. De Wekker, W. J. Shaw, R. L. Coulter, and K. J. Allwine, 2004: Observations of boundary-layer winds in an urban environment. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-25. De Wekker, S. F. J., L. K. Berg, K. J. Allwine, J. C. Doran, and W. J. Shaw, 2004: Boundary-layer structure upwind and downwind of Oklahoma City during the joint urban 2003 field study. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-26. Brown, M. J., D. Boswell, G. Streit, M. Nelson, T. McPherson, T. Hilton, E. R. Pardyjak, S. Pol, P. Ramamurthy, B. Hansen, P. Kastner-Klein, J. Clark, A. Moore, D. Walker, N. Felton, D. Strickland, D. Brook, M. Princevac, D. Zajic, R. Wayson, J. MacDonald, G. Fleming, and D. Storwold, 2004: Joint urban 2003 street canyon experiment. *Symp. on Planning, Nowcasting and Forecasting in the Urban Zone*, Seattle, WA, Amer. Meteor. Soc., January 12-16.

1-27. Brown, M. J., H. Khalsa, M. Nelson, and D. Boswell, 2004: Street Canyon flow patterns in a horizontal plane: measurements from the joint urban 2003 field experiment. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-28. Sol, S., P. Ramamurthy, E. R. Pardyjak, and J. C. Klewicki, 2004: Structure of turbulence in an urban street canyon. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-29. Pol, S. U. and M. J. Brown, 2005: Flow patterns at the ends of a street canyon: measurements from the joint urban 2003 field experiment. *Sixth Conference on*

the Urban Environment, Atlanta, Georgia, Amer. Meteor. Soc., January 30 - February 3.

1-30. Ramamurthy, P., S. Pol, E. Pardyjak, and J. Klewicki, 2004: Spatial and temporal variability of turbulent fluxes in the joint urban 2003 street canyon. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-31. Leach, M. J., J. H. Shinn, R. Leif, and G. Keating, 2004: Urban near-field dispersion. *Symp. on Planning, Nowcasting and Forecasting in the Urban Zone*, Seattle, WA, Amer. Meteor. Soc., January 12-16.

1-32. Grimmond, C. S. B., H. –B. Su, B. Offerle, B. Crawford, S. Scott, S. Zhong, and C. Clements, 2004: Variability of sensible heat fluxes in a suburban area of Oklahoma City. *Symp. on Planning, Nowcasting and Forecasting in the Urban Zone*, Seattle, WA, Amer. Meteor. Soc., January 12-16.

1-33. Zajic, D., H. J. S. Fernando, M. Princevac, and R. Calhoun, 2004: Flow and turbulence in urban canopies. *Fifth Conference on the Urban Environment,* Vancouver, British Colombia, Amer. Meteor. Soc., August 23-27.

1-34. https://ju2003-dpg.dpg.army.mil, website accessed on 27 March 2003.

1-35. Clawson, K. L. and D. Lacroix, 2004: Quality Assurance for the Time-Integrated Samplers, NOAA ARL FRD, frd_samplers_ReadMe.txt dated 6 February 2004.

1-36. Clawson, K. L., 2005: RE: JU2003 Questions …, e-mail communication from NOAA ARL FRD, 19 May 2005.

1-37. Guibas, L. J., D. E. Knuth, and M. Sharir, 1992: "Randomized Incremental Construction of Delaunay and Voronoi Diagrams," *Algorithmica* 7: 381-413, 1992. Also, see http://www.gris.uni-tuebingen.de/gris/proj/dt/dteng.html.

1-38. a) Warner S., N. Platt, and J. F. Heagy, 2005: Comparisons of transport and dispersion model predictions of the European tracer experiment: area-based and population-based measures of effectiveness. *Atmospheric Environment*, **39**, 4425-4437. b) Warner S., N. Platt, and J. F. Heagy, 2004: *Comparisons of Transport and Dispersion Model Predictions of the European Tracer Experiment: Area-Based and Population-Based Measures of Effectiveness,* IDA Paper P-3915, 139 pp. (Available electronically at DTIC STINET ada427807.)

1-39. Warner S., N. Platt, and J. F. Heagy, 2003: *Application of User-Oriented MOE to Transport and Dispersion Model Predictions of the European Tracer Experiment*, IDA Paper P-3829, November 2003. (Available electronically [DTIC STINET ada419433] or on CD via e-mail request to Steve Warner at swarner@ida.org or a mail request to Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882).

CHAPTER 2

PREPARATION OF JU03 PREDICITONS AND METHODOLOGIES FOR COMPARISON TO OBSERVED CONCENTRATIONS

2. PREPARTION OF JU03 PREDICITONS AND METHODOLOGIES FOR COMPARISON TO OBSERVED CONCENTRATIONS

This first section of this chapter provides a description of the preparation of Urban HPAC predictions of JU03 that were used for this study. This section includes a description of the meteorological inputs that were used to create these predictions and some discussion of initial analyses of meteorological data, some of which was not used for this current comparative study. The second part of this chapter describes the methodologies used to compare predictions to observations and sets of predictions to each other.

A. PREPARATION OF URBAN HPAC PREDICTIONS OF JU03

1. Transport and Dispersion Model Modes of Operation

For this study, we examined five types of Urban HPAC predictions: HPAC (v4.04 SP3 which included the vendor-provided MSS) with surface type entered as "urban," denoted baseline or "UC" (for urban canopy); HPAC with the UDM toggled on, denoted "DM"; HPAC with the UWM toggled on, denoted "WM"; and HPAC with both the UDM and the UWM toggled on, denoted "DW." We also examined predictions created with the newest urban feature, MSS, denoted as "MS." In general, default settings were used to create HPAC predictions. Our choices for several settings that were not necessarily set to the nominal default are described here. First, all predictions were created with the "Boundary Layer Points" setting fixed at 51 (higher resolution than default), a "Puff Split Grid Level" set at 2, and the choice "Ultimate" under the "Limits" tab. The choice of "Ultimate" sets the maximum grid cells per surface field to 100,000 and the maximum number of puffs to 60,000.

Predicted concentrations were output (typically), every 150 seconds, to a specified sampler file, which included the locations and altitudes of the samplers. Conditional averaging was set at 300 seconds, which was consistent with our expectations of comparing predicted concentrations to observed concentrations at a time resolution no higher than once per 5 minutes. The conditional averaging parameter in HPAC allows

one to define the averaging time for defining the diffusive component of turbulence and as such, is used to scale the velocity variances that determine puff diffusion. This choice of conditional averaging is consistent with HPAC technical manual guidance.

a. Preparation of UDM Predictions, DM

For the DM sets of predictions, UDM-unique modeling parameters were set at the defaults with "Channeling effects" turned on. UDM parameter defaults included "Building Interactions" set to Full; "Puff Splitting" and "Puff Merging" set to Medium; and "run Complexity" set to Full. Figure 2-1 compares (as an example) predicted surface dosages after the three releases of IOP 10 for UC- and DM-based predictions. The metrological inputs for these example predictions (and others) will be discussed later. The smoother, broader features associated with the DM predictions relative to the UC predictions can be discerned from these surface dosage plots.

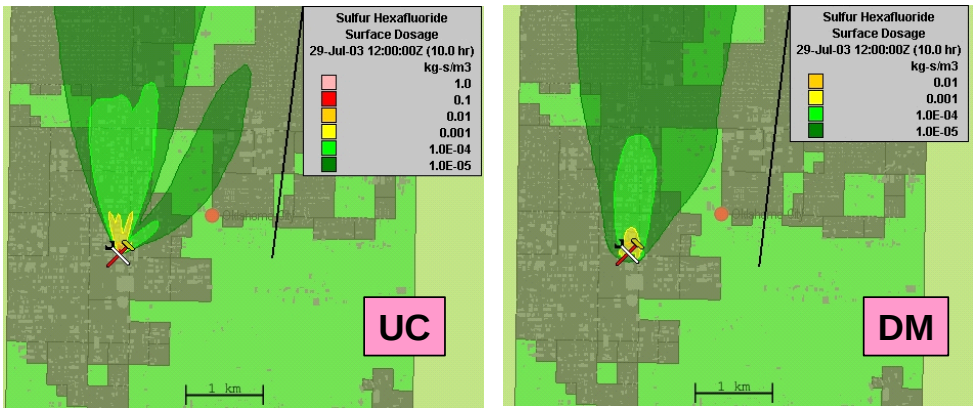


Figure 2-1. Predicted Surface Dosage After the Three IOP 10 Releases

These UC and DM predictions were created using the baseline meteorological input option, which is described later in this section.

b. Preparation of UWM-Based Predictions, WM and DW

For the WM sets of predictions, the UWM-unique modeling parameters were set as shown in Table 2-1. Note that for the DW sets of predictions, the UDM setting "Channeling effects" was turned off. This was done to avoid the possibility of essentially "double-counting" an effect when both UDM and UWM are toggled on since UWM is designed to already account for such effects. Land cover type (e.g., barren, shrubland, shrubland/grassland) was included in all DM, WM, and DW predictions by including a terrain file and enabling the "land cover" feature.

Figure 2-2 shows a "screen dump" from a UWM-based prediction as well as the contents of the uwm.dat file that was used for these predictions. Figure 2-3 provides an example of the UWM-generated wind vectors over the CBD using the parameter settings shown in Table 2-1 and Figure 2-2 for IOP 4 Release 1.

Table 2-1. UWM-Unique Parameter Values[1]

Parameter	Value
Resolution	Fine
Domain	Automatically calculated
Minimum horizontal resolution	5 m
Maximum domain size	5000 m
Minimum domain size	2500 m

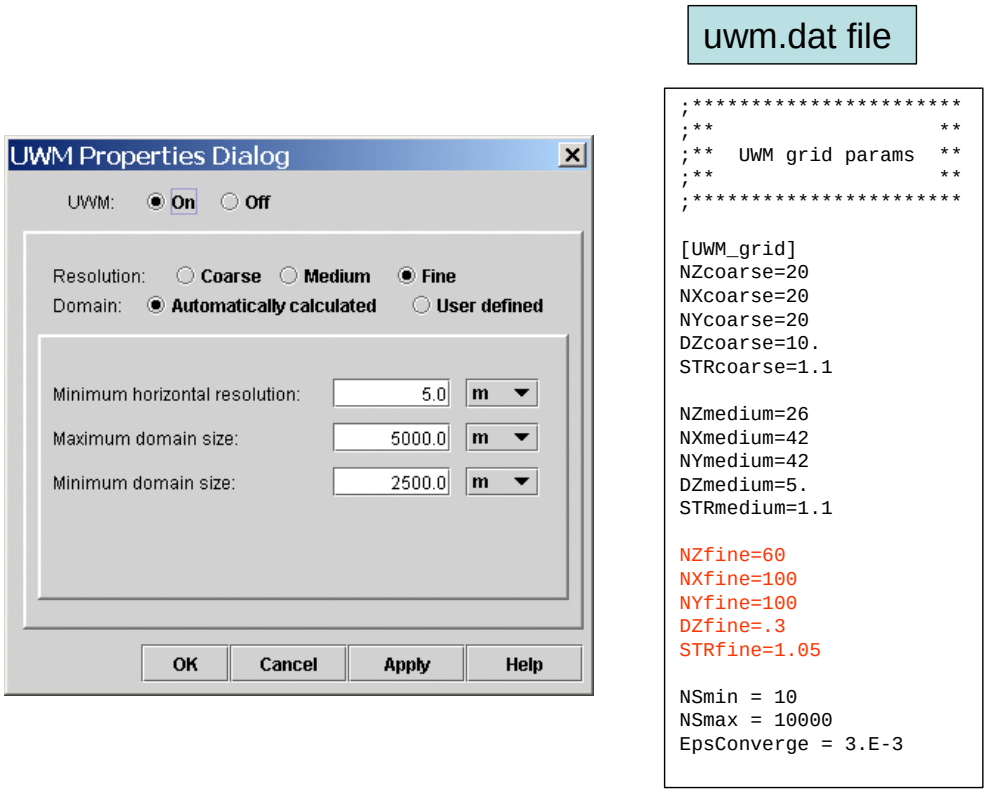


Figure 2-2. Screen Dump of Key UWM Parameter Settings and Contents of the uwm.dat File

---

1 The parameters DZfine and NZfine in the uwm.dat file were reset from the defaults of 5 and 36 to 0.3 and 60, respectively allowing for greater vertical resolution and 60 layers. In addition, the parameters NXfine and NYfine were changed from 66 to 100 to ensure that a horizontal resolution of about 10 m could be obtained.

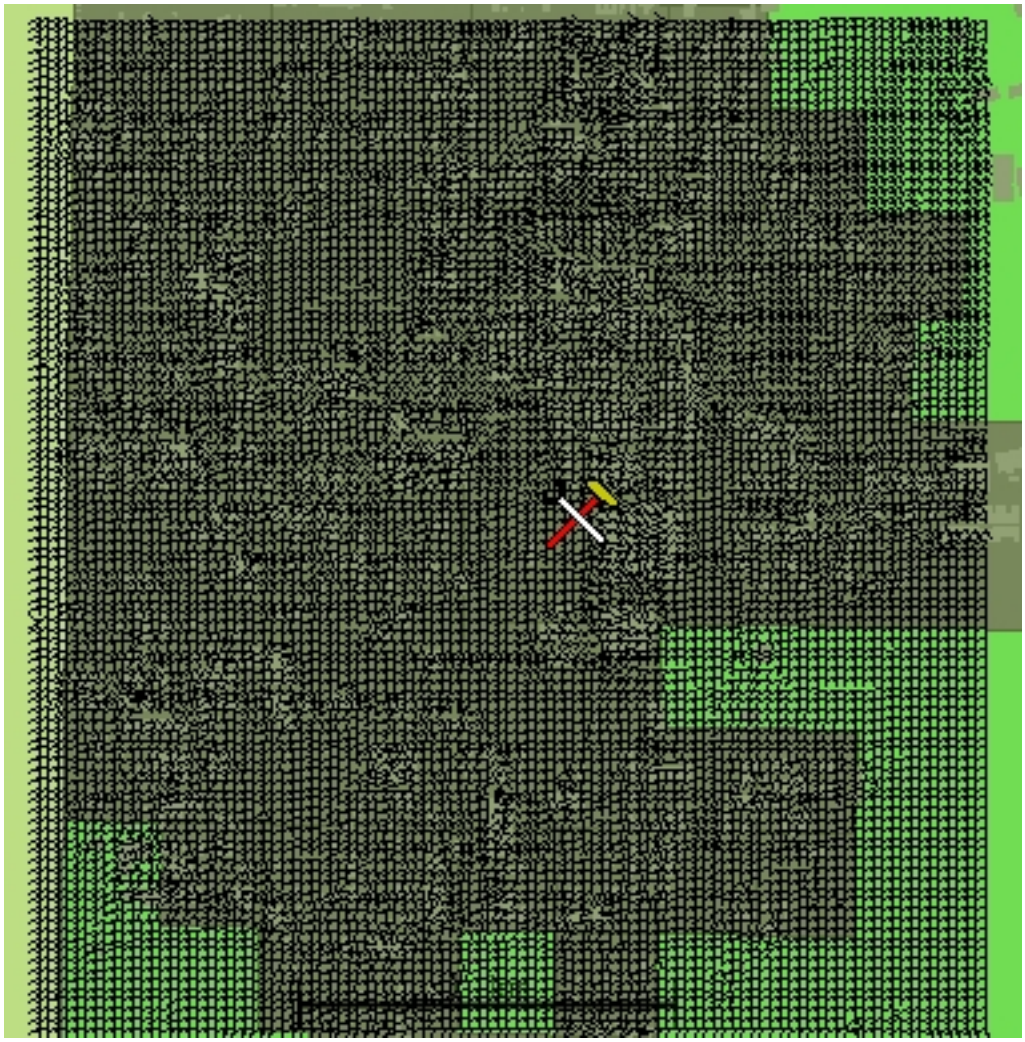


Figure 2-3. UWM-Generated Wind Vectors in the CBD for IOP 4

View is for 5 minutes after the 1st release at 1 meter AGL.

c. Preparation of MSS Predictions, MS

For this study, we created predictions of JU03 using the newly available MSS. MSS was delivered to IDA as a replacement set of Dynamic Link Libraries (DLLs) that needed to be installed over existing (sometimes just placeholder) DLLs in the regular HPAC 4.04 Service Pack 3 release. Some limitations exist in the present MSS configuration; they affected our JU03 protocol as follows: (1) only a single release per HPAC project is supported and (2) the HPAC default domain must be in the Universal Transverse Mercator (UTM) coordinate system. We solved the first of these limitations by redefining JU03 continuous releases in terms of a single HPAC project per release instead of a single HPAC project per IOP (as was the case for the other Urban HPAC

configurations). The second limitation involving MSS required the SCIPUFF project domain to be in the UTM coordinate system. This led to a more elaborate procedure. For meteorological input options (to be discussed in the next section) that used SWIFT to preprocess the meteorological data, this requirement was satisfied by default since SWIFT requires UTM coordinates. For meteorological input options that used MC-SCIPUFF to preprocess the meteorological data, we exploited a SWIFT limitation that allowed us to cause SWIFT to abort and HPAC to switch to MC-SCIPUFF while keeping the SCIPUFF default domain in UTM coordinates.

User-changeable Micro SWIFT input parameters include "Horizontal grid resolution," "Horizontal domain size," "Vertical grid resolution," and "Vertical Grid Clustering." The actual values for the first three parameters are hard-coded in Micro SWIFT, but the user is allowed to select from a set of three (predefined) values. By default, Micro SWIFT set these parameters to the "middle" values. We used "upper" (i.e., higher resolution) values whenever possible resulting in a "Horizontal grid resolution" of 3 meters, a "Horizontal domain size" of 1 km, and a "Vertical grid resolution" of 31 points. The "Vertical Grid Clustering" was kept at the default value preset by the Micro SWIFT code.

User-changeable Micro SPRAY parameters include "Maximum number of particles" and "Maximum height above ground for viewing concentration data." The default value for the maximum number of particles is 50,000 particles, and we doubled it to 100,000 for our runs and kept the second parameter at the default value provided by MSS. Key MSS parameters are highlighted in Figure 2-4. For MSS-based predictions, concentrations were output every 60 seconds (because MSS currently only outputs 60-second average concentrations).

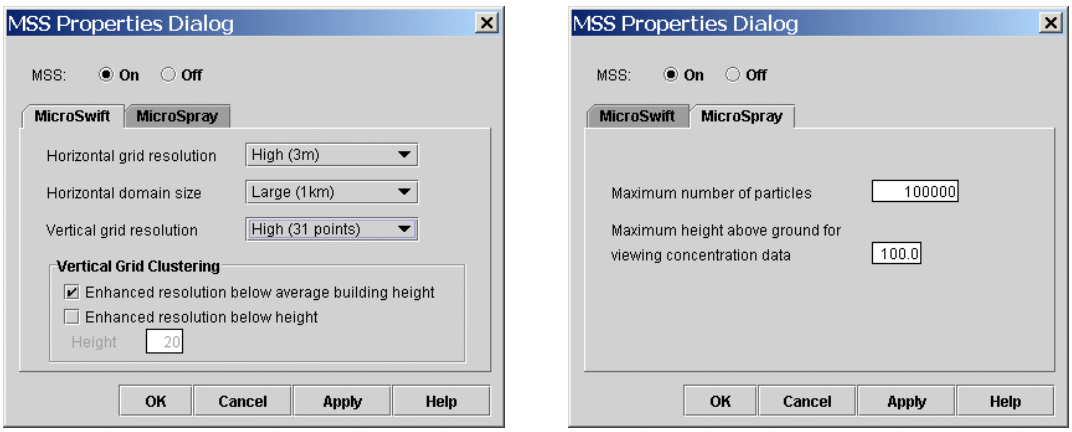
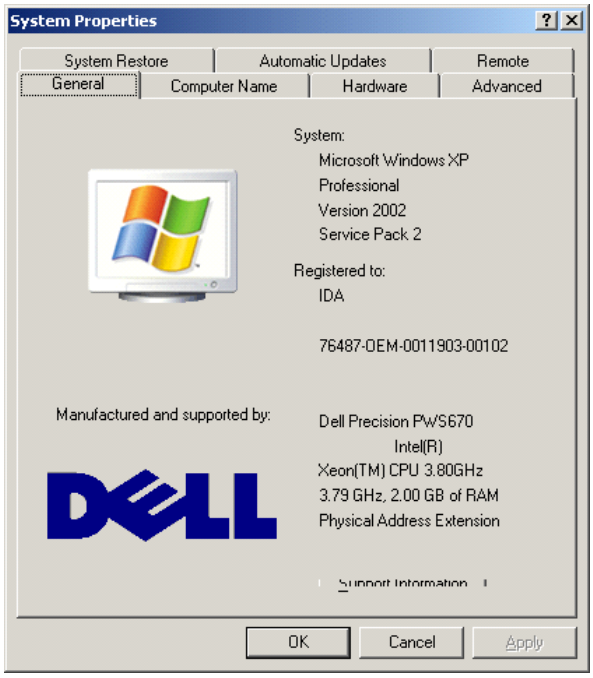


Figure 2-4. Screen Dump of Key MSS Parameter Settings

d. Examination of Urban HPAC Run Times

As part of this study, the time to complete a model run was measured for several Urban HPAC configurations. In addition to the UC, DM, WM, DW, and MS sets of predictions described previously, we created lower-resolution, faster-running configurations denoted WM_{Lo}, DW_{Lo}, and MS_{Lo}. These predictions were created on a Dell Precision PWS670 (3.80 GHz), Microsoft Windows XP Professional (Version 2002, Service Pack 2) as shown in Figure 2-5.

Figures 2-6 and 2-7 show screen dumps associated with the WM_{Lo} (and DW_{Lo}) and MS_{Lo} predictions, respectively. For WM_{Lo} and DW_{Lo}, the minimum horizontal resolution is set to 50 meters (similar to the default value) as opposed to 5 meters as was the case for the nominal (for this study) WM and DW predictions. Comparing the MS (Figure 2-4) and MS_{Lo} (Figure 2-7) parameter settings, shows several changes from MS to MS_{Lo} as follows: (1) horizontal grid resolution [High (3 meters) → Low (5 meters)], (2) horizontal domain size [Low (1 km) → Medium (0.8 km)], vertical grid resolution [High (31 points) → Low (21 points)], and maximum number of particles [100,000 → 50,000].



Double Processor, but only single CPU is utilized during timing runs

Figure 2-5. Screen Dump of Computer System Specifications

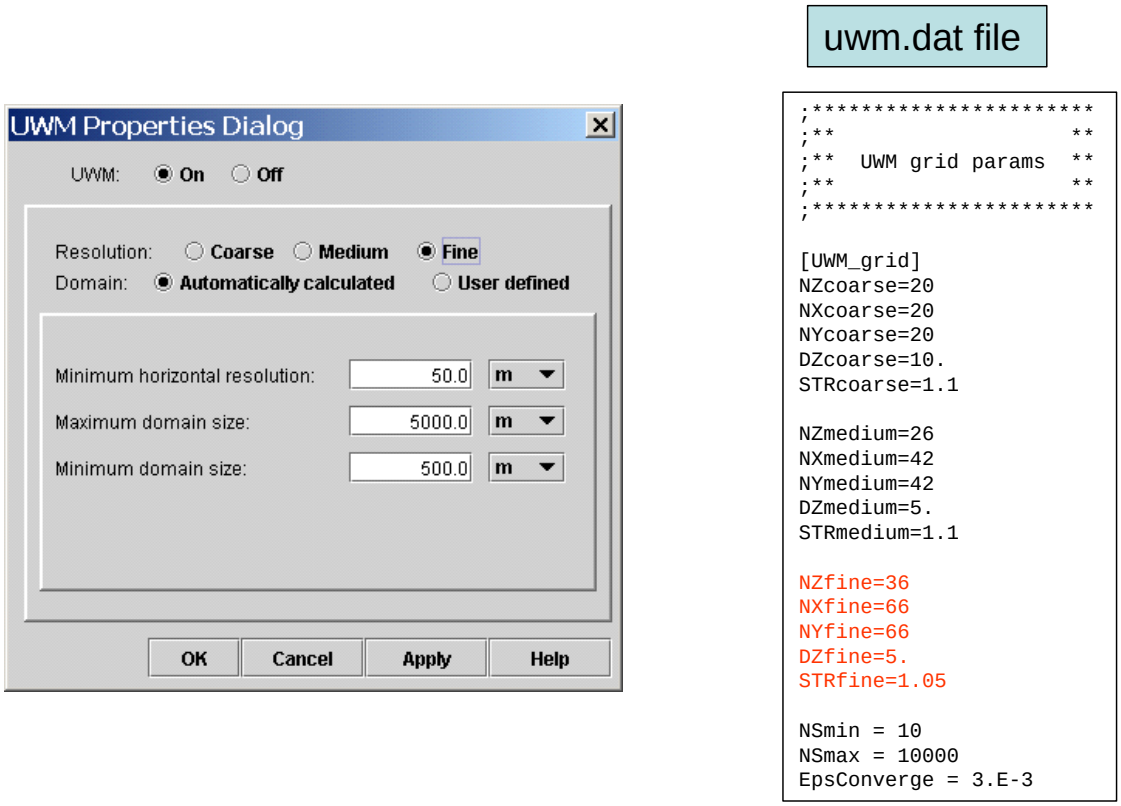


Figure 2-6. Screen Dump of Key UWM Parameter Settings and Contents of the uwm.dat File for the Lower-Resolution WM and DW Predictions – WM_{Lo} and DW_{Lo}

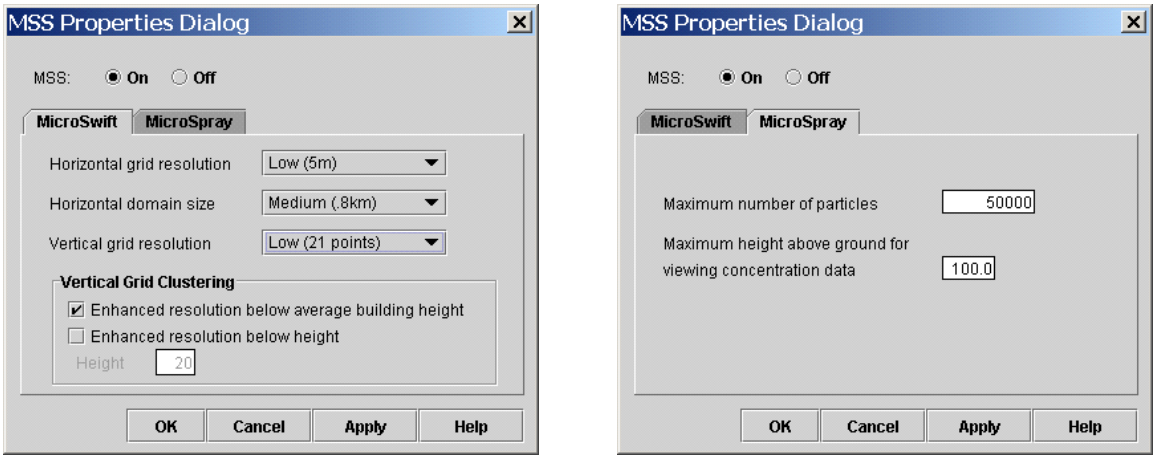


Figure 2-7. Screen Dump of Key MS_{Lo} Parameter Settings

Figures 2-8 and 2-9 display the run times for each of the 29 releases (when run individually) and for each of the Urban HPAC configurations. Two meteorological input options were used to generate these predictions: BAS (baseline, Figure 2-8) and PO7 (post office roof, Figure 2-9), which are used only as examples here. Meteorological input options will be discussed in the next section of this chapter.

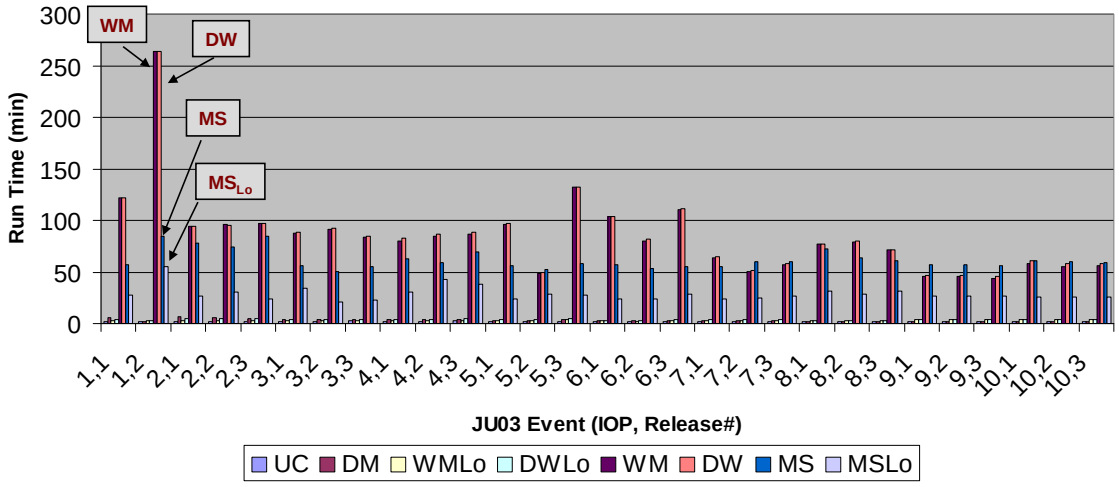


Figure 2-8. Run Times for Urban HPAC Predictions of 29 JU03 Releases Using Various Configurations and the BAS Meteorological Input Option

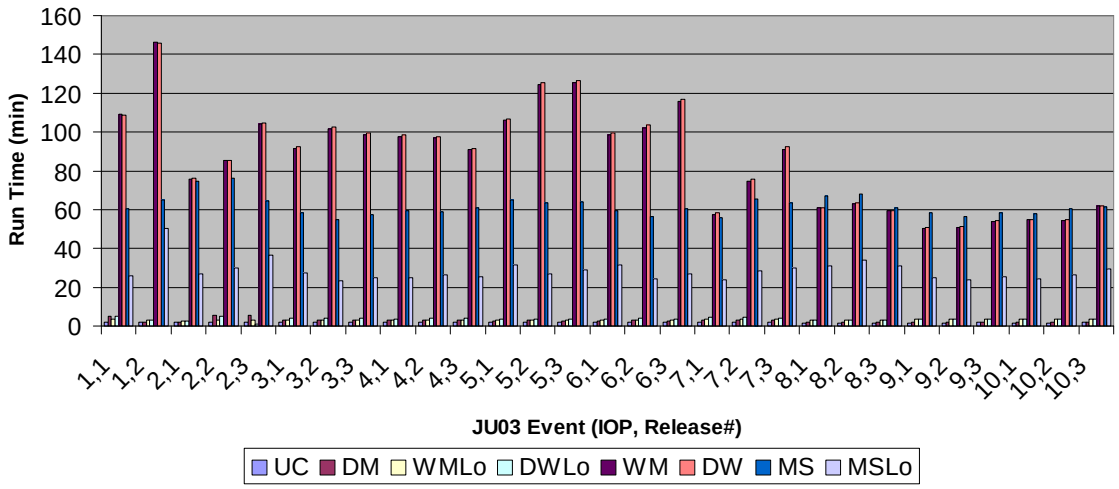


Figure 2-9. Run Times for Urban HPAC Predictions of 29 JU03 Releases Using Various Configurations and the PO7 Meteorological Input Option

Table 2-2 summarizes some run time statistics – median, average, minimum, and maximum – associated with these predictions of 29 releases. Figure 2-10 plots the median values. It is seen that the WM, DW, MS, MS_{Lo} configurations require substantially more run time than the UC, DM, WM_{Lo}, and DW_{Lo} configurations. For perspective, we note that the Joint Effects Model (JEM) program includes SCIPUFF and is considering the inclusion of urban atmospheric and dispersion code as an upgrade. JEM requires that a single release be predicted in about 5 minutes [Ref. 2-1]. Specifically, the JEM operational requirements document states, "JEM, running without

advanced features turned on (i.e., secondary evaporation, complex terrain, microscale meteorology), shall provide hazard prediction data and graphical display, for up to two known (e.g., location, agent, dissemination) source terms within 10 minutes (8 minutes [O]).” Figure 2-11 shows that the lower resolution and non-MSS Urban HPAC configurations meet this requirement.

Table 2-2. Run Time Statistics for Predictions Using Various Urban HPAC Configurations: BAS and PO7 Meteorological Input Options

Time (min)	UC	DM	WM	WM_{Lo}	DW	DW_{Lo}	MS	MS_{Lo}
				BAS				
Median	2.1	3.1	80.5	3.0	82.4	3.7	59.3	26.5
Average	2.1	3.3	85.0	3.0	85.9	3.8	61.7	28.7
Minimum	1.9	2.1	44.0	2.4	46.0	2.7	50.1	21.3
Maximum	2.5	6.2	263.6	3.6	263.4	5.2	85.1	55.1
				PO7				
Median	1.9	2.8	91.1	3.2	92.2	3.8	60.6	26.9
Average	1.9	2.9	86.3	3.2	86.8	3.7	61.8	28.4
Minimum	1.7	1.9	50.3	2.8	51.0	2.8	54.6	23.5
Maximum	2.0	5.7	146.5	3.7	145.8	5.1	76.1	50.4

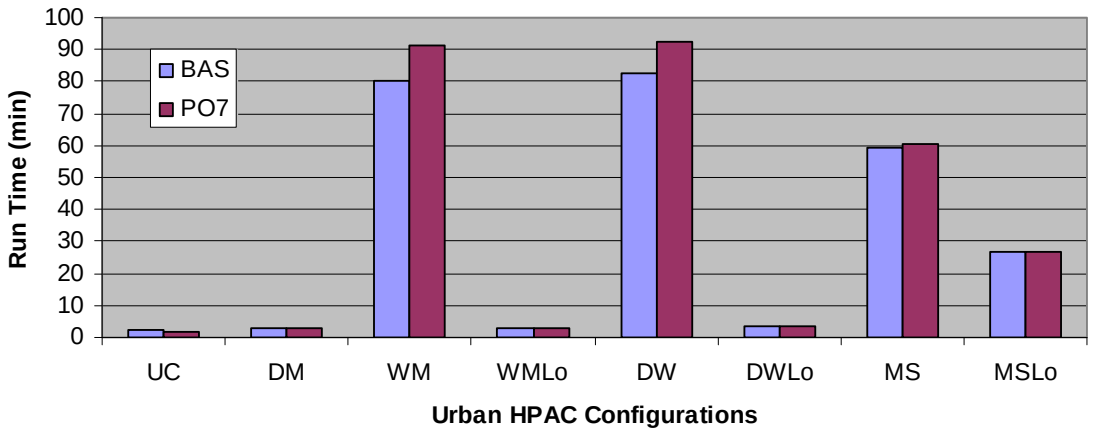


Figure 2-10. Median Run Times (29 JU03 Releases) for Eight Urban HPAC Configurations and Two Meteorological Input Options

As a quick side analysis using hypothesis test procedures described later in this chapter, we compared predictions associated with the low and high resolution versions of WM, DW, and MS. We did not find substantial, significant differences due to the lower resolution configurations when compared with their higher resolution pairs. In fact, only

very limited evidence suggested that the higher resolution predictions ever corresponded to significant improvements over the lower resolution predictions. That is, the overall fits to the observations were quite similar across the range of resolutions that were examined. However, this quick "side" analysis was not the main subject of this study. Although no substantial significant differences in prediction quality were observed as a function of model resolution, Urban HPAC predictions for the rest of this study were run using only the higher resolution modes of UWM (i.e., WM and DW) and MSS (i.e., MS) essentially, in order to give the benefit of the doubt.

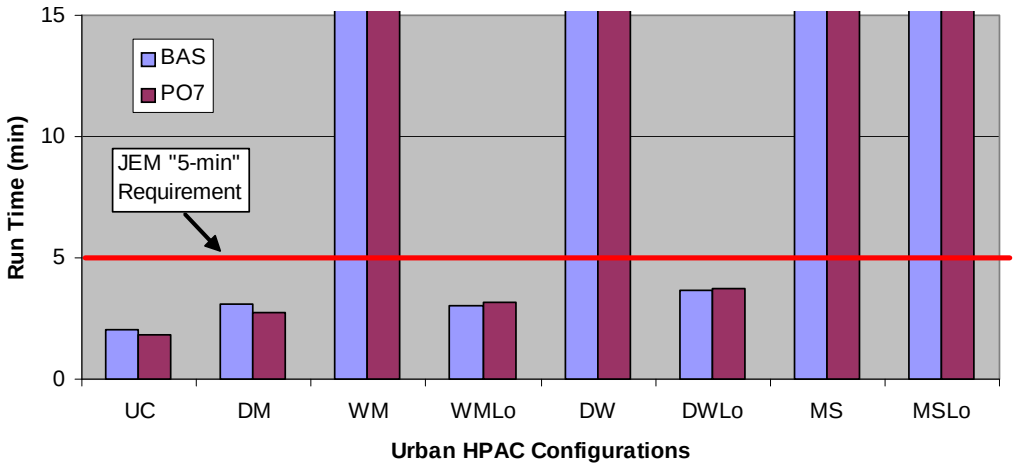


Figure 2-11. Comparison of JEM Run Time Requirement and Median Run Times (29 JU03 Releases) for Eight Urban HPAC Configurations and Two Meteorological Input Options

2. Meteorological Input Options Used for this Comparative Study

A large variety of meteorological measurements were collected during JU03 [Ref. 2-2]. The main goal of this paper is to describe the comparative results of JU03 predictions created with varying Urban HPAC configurations – UC, WM, DM, DW, and MS. Several MET options were examined as a part of this study. This section provides a brief description of the five input options that were selected for this comparative study and discusses, briefly, some of the options that were rejected. Although some discussion of the effect of MET options on predictions is described in the results section of this paper, a follow-on study is planned to further examine the impact of varying MET options.

The shorthand notations for the five MET options that were chosen for this study are (1) BAS, (2) BRB, (3) PNA, (4) ACA, and (5) PO7. Each MET option will be described below. A final section will describe some of the other MET options that were examined but not used for this final study.

a. BAS

The BAS MET option was designed to correspond to a baseline situation where the meteorological information is consistent with what could have been retrieved from the DTRA meteorological server at some point (~ 2 hours or more) after the release. That is, this information corresponds to *assimilated observations during the release,* as opposed to forecast information. Meteorological sources within 30 km of Oklahoma City were considered for BAS. Surface wind speed and direction observations from stations between 12 and 28 km and upper air (wind speed and direction) observations from a station 28 km southeast were used as is shown in Figure 2-12. Surface observations were typically updated every hour and the upper air measurements were collected twice daily. The diagnostic wind field model, SWIFT, resident within Urban HPAC, was used to create gridded wind fields from the BAS input meteorological information. SWIFT was run with an update interval of 15 minutes.

- **Surface**
  - Source: University of Utah Mesonet (MesoWest)
    - » Stations: KOUN (Norman, OK), KOKC, KPWA, KTIK (Airport)
- **Upper Air**
  - Source: University of Wyoming
  - Station: KOUN (Norman, OK)

Red – Surface
Blue – Profile
Black - Releases

The prevailing wind is from South

Figure 2-12. Locations of BAS MET Option Meteorological Observations

Comparisons of Urban HPAC predictions created with the BAS MET option generally showed good agreement with the observations in terms of plume direction. As shown in Figure 2-13, a significant plume direction discrepancy was discovered for IOP 1, Release 2. This discrepancy was further investigated.

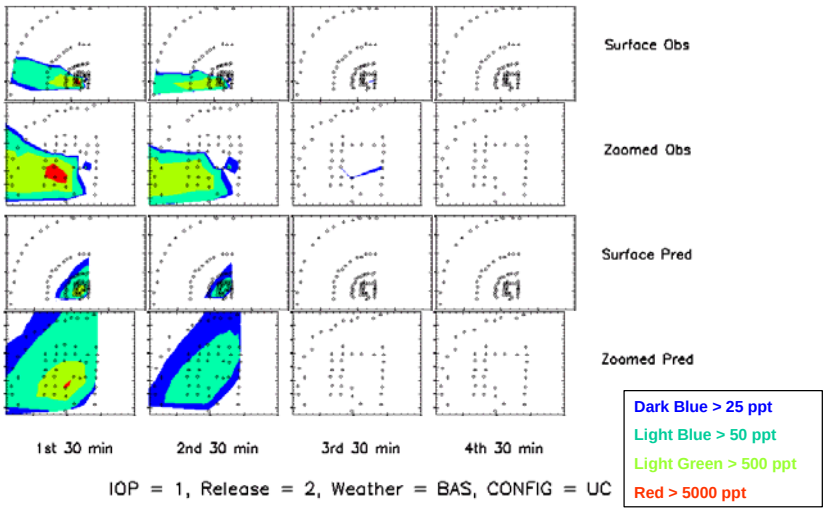


Figure 2-13. Illustration of Plume Direction Discrepancy for Release 2 of IOP 1 When Using the BAS MET Option and UC Urban HPAC Mode ("BAS_UC")[2]

Two additional separate predictions of this release were created using (1) only the surface observations from the BAS option and (2) only the upper air profile observations. Figure 2-14 presents the resulting predicted plumes (captured from the HPAC screen). It appears that the surface observations lead to predictions that generally are in agreement with the concentration observations. It also appears that the upper air profile observations may be causing the plume direction discrepancy associated with IOP 1, Release 2.

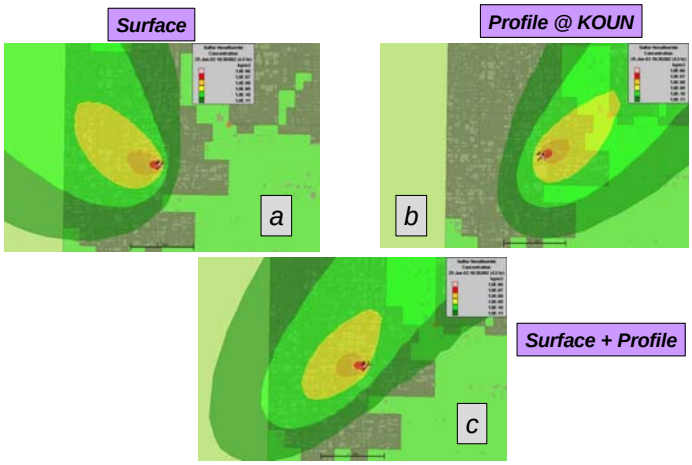


Figure 2-14. Illustration of Plume Direction Differences for Release 2 of IOP 1 When Using the BAS MET Options (a) Surface-Only Observations, (b) Upper Air Profile Observations, and (c) Surface and Upper Air Profile Observations

---

2 Figure 2-13 is of similar format to Figure 1-12. Additional details of this format can be found in the caption associated with Figure 1-12.

Figures 2-15 and 2-16 show portions of tables that list the actual observations associated with the surface and upper air profile observations, respectively. Note that the time of release for IOP 1, Release 2 was 18:00 UTC. First, the surface observations (Figure 2-15) indicate a significant change in wind direction starting near 18:00 UTC. The winds move from the south (160-180 degrees) to the east (80-90 degrees, shown in red font).[3] The observed plume (Figure 2-13) clearly moves in a direction consistent with flow from the east. Figure 2-16 indicates that upper air profile measurements were collected at 12:00 UTC and 0:00 UTC, with the actual release right in the middle at 18:00 UTC. The observed upper air wind direction is seen to shift considerably between 12:00 UTC and 0:00 UTC. It appears that Urban HPAC's use of the sparse upper air wind direction data associated with the BAS MET option resulted in an assimilated wind field that was too heavily influenced by upper air measurements and ultimately led to the identified plume direction discrepancy for IOP 1, Release 2. One might hypothesize that real-time meteorological observations might alleviate this specific plume discrepancy problem. Three of the MET options that we consider below include real-time meteorological observations – PNA, ACA, and PO7 – and we will briefly consider this BAS option discrepancy later in this section.

b. BRB

The BRB MET option was designed to correspond to a baseline situation in which a gridded numerical weather assimilation was used as input to Urban HPAC. The BRB MET option corresponded to a GCAT prediction of vertical profiles (i.e., wind velocity as a function of height AGL) at many grid locations – that is, an HPAC profile ".prf" file was used.[4] These files can be thought of as surrogates for "gridded" numerical weather assimilations that could be available on the DTRA meteorological server several hours after an event.

GCAT was developed by the National Center for Atmospheric Research (NCAR) and NGIC. GCAT combines the 5th version of the Pennsylvania State University/NCAR (regional atmospheric) Mesoscale Model (MM5) with the Four Dimensional Data

---

[3] The "-9999" value in the wind direction column in Figure 2-15 denotes an unreliable or missing measurement.

[4] We thank the National Ground Intelligence Center, Charlottesville (NGIC), VA, for providing the GCAT-based gridded profiles for JU03. The shorthand notation "BRB" stands for "Baseline" followed by the initials, "RB," of the NGIC scientist that made these profiles available to IDA.

#	CREATOR:	Weather	File	Editor	Version	1.17	)c/129.246.70.141)				
#	DATE:	6/29/2005	17:34								
#	SOURCE:	OBS									
#	EDITED:	YES									
#	EFERENCI	AGL									
#	TYPE:	SERVATION									
#	EREFEREN	UTC									
#	MODE:	OBS	ALL								
SURFACE											
12											
ID	YYYYMMDI	HOUR	LAT	LON	ELEV	Z	WDIR	WSPD	P	T	HUMID
		HOURS	N	E	M	M	DEG	M/S	MB	C	%
-9999											
KTIK	20030629	15.92	35.4167	-97.3833	394	10	200	4.47	973	26	54
KOUN	20030629	16.17	35.2167	-97.45	357	10	200	4.47	978	27	70
KOUN	20030629	16.58	35.2167	-97.45	357	10	200	4.47	978	28	62
KOKC	20030629	16.83	35.3886	-97.6003	394	10	**150**	4.47	972	26	61
KOUN	20030629	16.83	35.2167	-97.45	357	10	**170**	4.47	977	28	55
KPWA	20030629	16.83	35.5411	-97.6467	396	10	**150**	4.47	972	26	61
KTIK	20030629	16.92	35.4167	-97.3833	394	10	**180**	4.02	972	27	48
KOUN	20030629	17.17	35.2167	-97.45	357	10	**150**	3.58	977	28	55
KOUN	20030629	17.5	35.2167	-97.45	357	10	**150**	4.02	977	28	55
KOKC	20030629	17.83	35.3886	-97.6003	394	10	**100**	4.47	972	26	58
KOUN	20030629	17.83	35.2167	-97.45	357	10	**140**	3.58	977	28	58
KPWA	20030629	17.83	35.5411	-97.6467	396	10	**110**	3.58	972	26	58
KTIK	20030629	17.92	35.4167	-97.3833	394	10	**140**	2.68	972	26	54
KOUN	20030629	18.17	35.2167	-97.45	357	10	110	5.81	976	29	55
KOUN	20030629	18.5	35.2167	-97.45	357	10	**80**	5.36	976	29	58
KOKC	20030629	18.83	35.3886	-97.6003	394	10	**80**	4.47	972	27	58
KOUN	20030629	18.83	35.2167	-97.45	357	10	**90**	5.81	976	29	62
KPWA	20030629	18.83	35.5411	-97.6467	396	10	**50**	3.58	972	26	58
KTIK	20030629	18.92	35.4167	-97.3833	394	10	**-9999**	1.34	972	26	58
KOUN	20030629	19.17	35.2167	-97.45	357	10	**90**	3.13	977	29	62
KOUN	20030629	19.5	35.2167	-97.45	357	10	**-9999**	-9999	977	30	59
KOKC	20030629	19.83	35.3886	-97.6003	394	10	**0**	0	973	27	58
KOUN	20030629	19.83	35.2167	-97.45	357	10	**30**	1.34	977	30	59
KPWA	20030629	19.83	35.5411	-97.6467	396	10	**-9999**	-9999	973	27	54
KTIK	20030629	19.92	35.4167	-97.3833	394	10	**-9999**	1.34	973	27	51
KOUN	20030629	20.17	35.2167	-97.45	357	10	**50**	2.68	977	30	59

Figure 2-15. Relevant Excerpts from BAS MET Option Surface Station Files for IOP 1, Release 2

ID:	72357	20030629	**12**	35.2167	-97.45		ID:	72357	20030630	**0**	35.2167	-97.45
Altitude	DIR	SPEED					Altitude	DIR	SPEED			
2	**240**	1.543	977	19.8	85		2	**70**	3.601	975	28.4	66
255	**225**	8.746	949	20.6	75		255	**75**	3.087	948	25.9	68
325	**225**	9.26	941	20.8	73		468	**90**	3.087	925	23.8	69
380	**225**	9.774	935	20.8	88		559	**95**	3.087	915	22.9	72
473	**225**	10.29	925	20.6	82		864	**110**	3.601	884	20	83
559	**225**	9.774	916	20.2	72		1201	**140**	3.087	850	16.8	97
643	**229**	9.774	907	19.8	64		1474	**180**	4.116	823	14.4	88
864	**240**	10.29	884	18	77		1517	**186**	4.116	819	14	87
1110	**255**	8.746	859	16	94		1590	**195**	3.601	812	13.8	100
1200	**260**	8.231	850	16.4	85		1715	**212**	3.087	800	12.8	73
1475	**274**	6.173	823	15.2	92		1779	**220**	3.087	794	13.6	62
1652	**283**	4.63	806	15.2	72		1832	**235**	2.572	789	14.2	55
1779	**290**	3.601	794	14.2	73		2083	**310**	1.543	766	12.5	57
2083	**275**	4.116	766	11.8	75		2388	**340**	3.601	738	10.4	60

Note that release time is at 18:00 (half way between these two profiles)

Figure 2-16. Relevant Excerpts from BAS MET Upper Air Profile Observations Measured at KOUN, Norman, OK (28 km away from the Oklahoma City CBD)

Assimilation (FDDA) technique to produce fine-scale climatology analyses anywhere in the world [Ref. 2-3]. FDDA is a Newtonian relaxation-based continuous data assimilation technique, which allows the sequential insertion of observations into a continuously running mesoscale model with tuned temporal and spatial weights according to observation times and locations. This approach incorporates observations that are available at irregular times and locations.

For BRB, GCAT was run with 36 vertical levels, with the first level at 20 meters AGL, and a horizontal resolution of 3.3 km. SWIFT was used to create gridded wind fields from the BRB input meteorological information. SWIFT was run with an update interval of 15 minutes.

Figure 2-17 shows an example comparison of Urban HPAC predictions to JU03, IOP 7, Release 1 observations using the BRB MET option and MS. Figure 2-18 shows another BRB-based set of comparisons, this time for all five Urban HPAC modes and for IOP 3, Release 3.

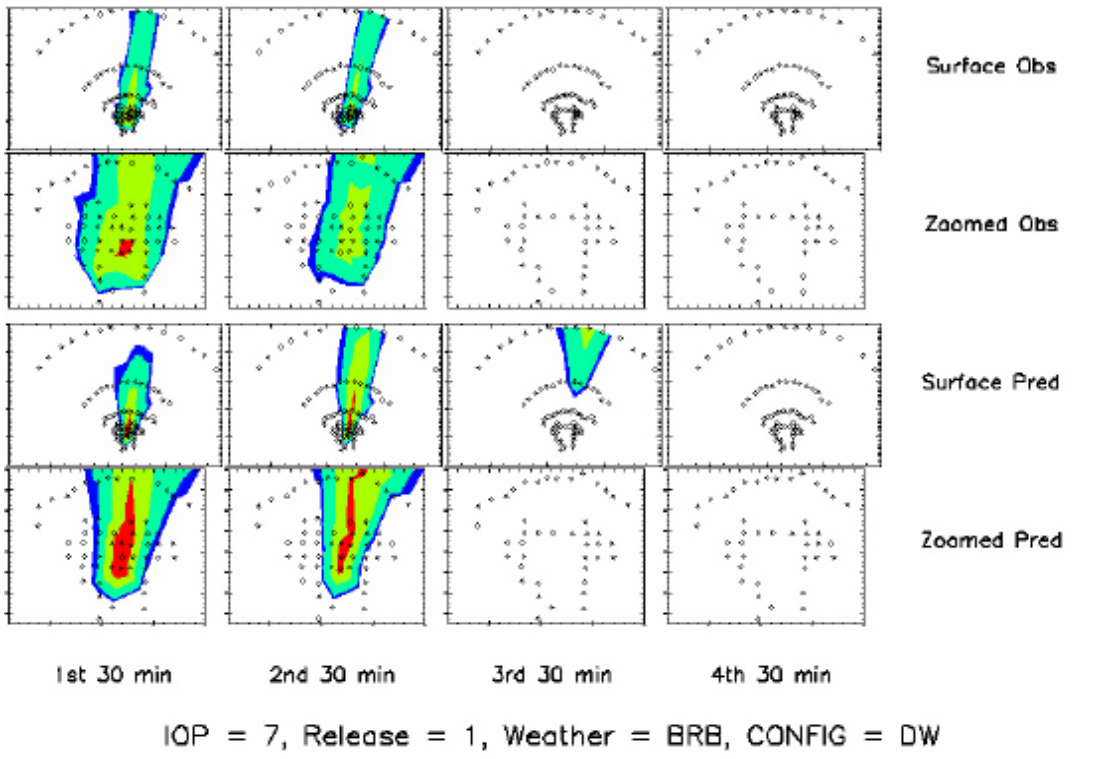


Figure 2-17. Comparisons of 30-Minute Average Concentration Contours for IOP 7, Release 1: Predictions Based on the BRB MET Option and the MS Urban HPAC Mode – "BRB_MS"

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

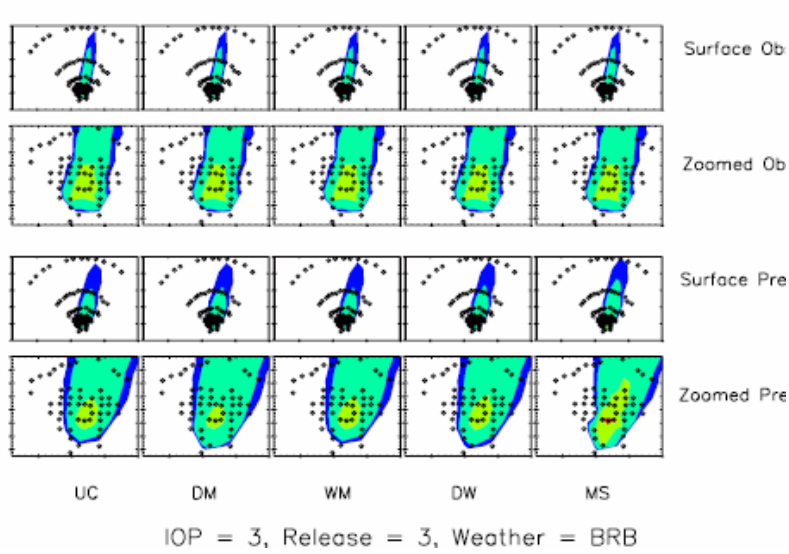


Figure 2-18. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 3, Release 3: All Predictions Based on the BRB MET Option

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

c. PNA

A large amount and variety of meteorological measurements were made in the vicinity of Oklahoma City during the JU03 field experiment. Figures 2-19 through 2-24 show the locations of some of these measurements. Table 2-3 provides definitions for the acronyms and shorthand notations used in these figures.

Table 2-3. Acronyms and Shorthand Notation Used in Figures 2-19 through 2-24

Acronym	Description
ANL	Argonne National Laboratory
BG	Botanical Gardens
CC	Christian Church
DPG	Dugway Proving Ground
FRD	Field Research Division (of NOAA ARL)
PNNL	Pacific Northwest National Laboratory
PWIDS	Portable Weather Information and Display Systems
RASS	Radio Acoustic Sounding System
SODAR	Sonic Detection and Ranging or simply Acoustic Sounder
U	University
UoU	University of Utah
WindTracer	LiDAR (Light Detection and Ranging) used for measuring radial wind velocities

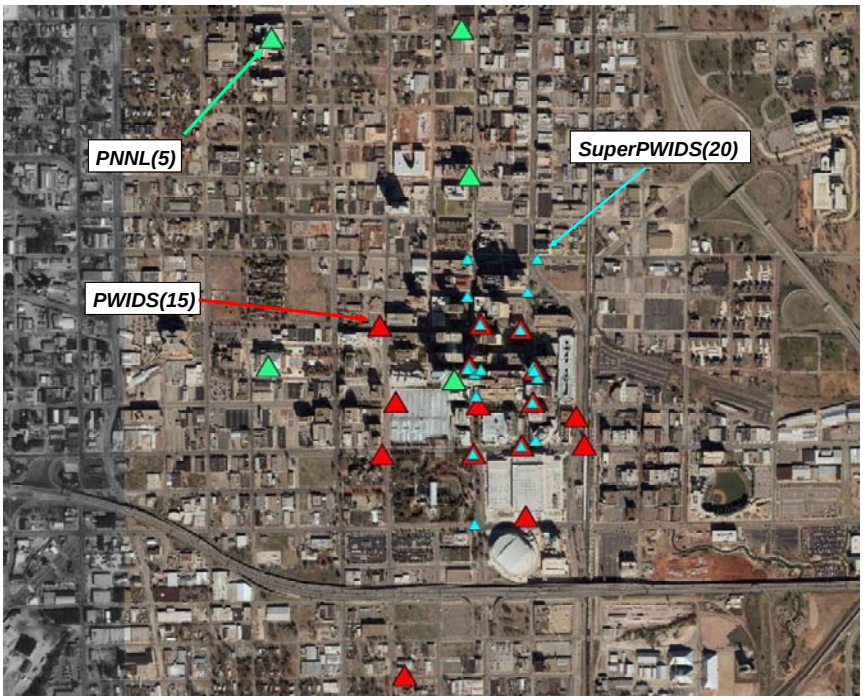


Figure 2-19. Locations of Surface Meteorological Stations in Oklahoma City for JU03

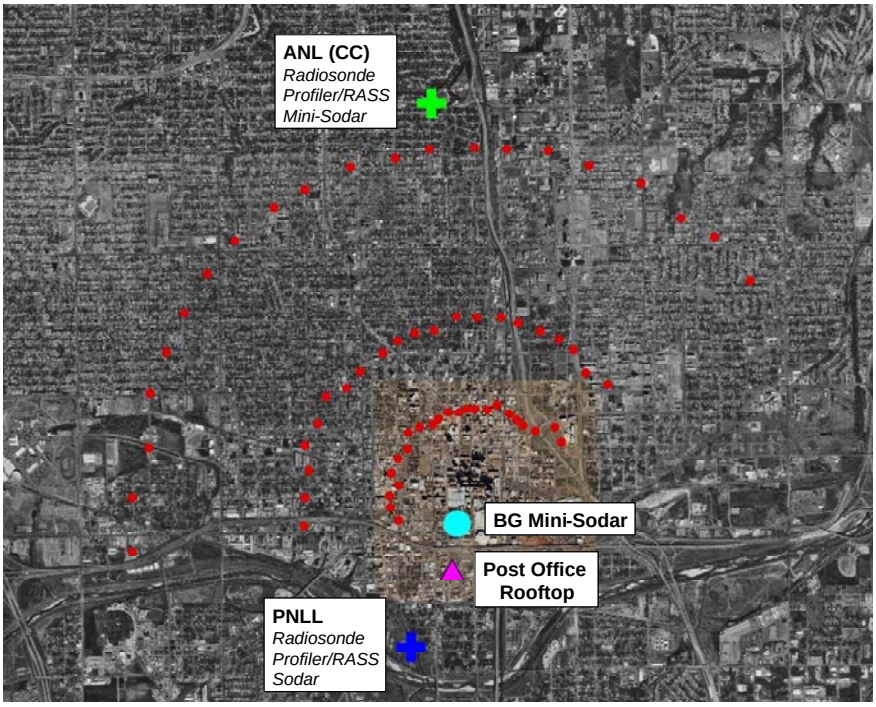


Figure 2-20. Locations of Some Upwind–PNNL, Downwind–ANL(CC), and Downtown–BG Vertical Wind Profile Measurements During JU03 and Location of the Post Office Rooftop Measurement Site

Red circles correspond to surface sampler locations.

Figure 2-21. Close-up View of the Downtown–BG Vertical Wind Profile Measurement Location

Blue pentagons correspond to IOP release locations.

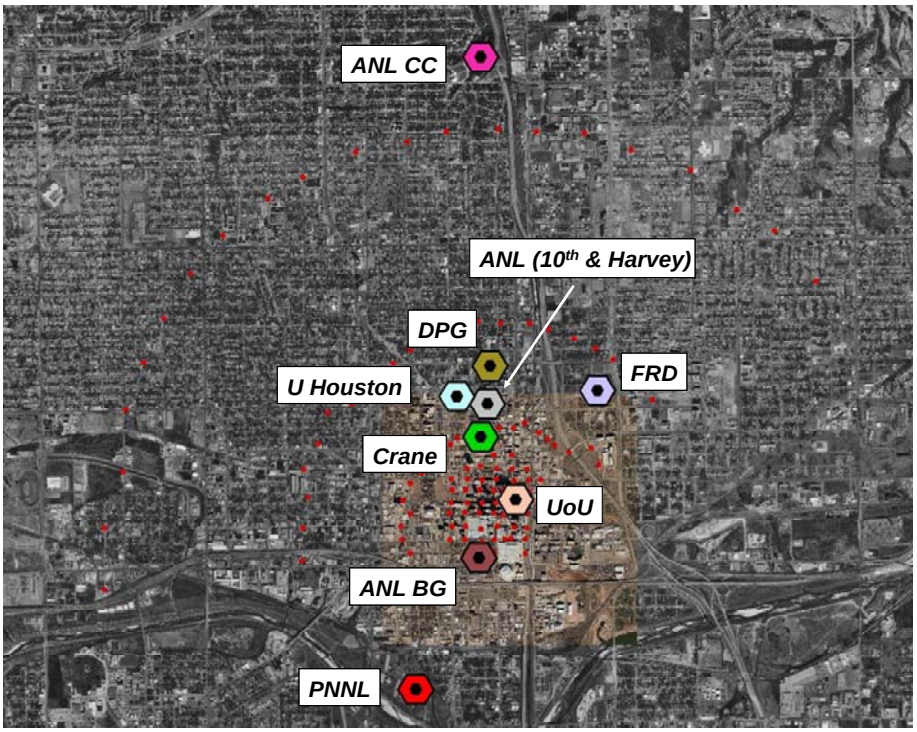


Figure 2-22. Locations of Some Vertical Wind Profile Measurements During JU03 Using SODARs, mini-SODARs, and Sensors on a Crane

Red circles correspond to surface sampler locations.

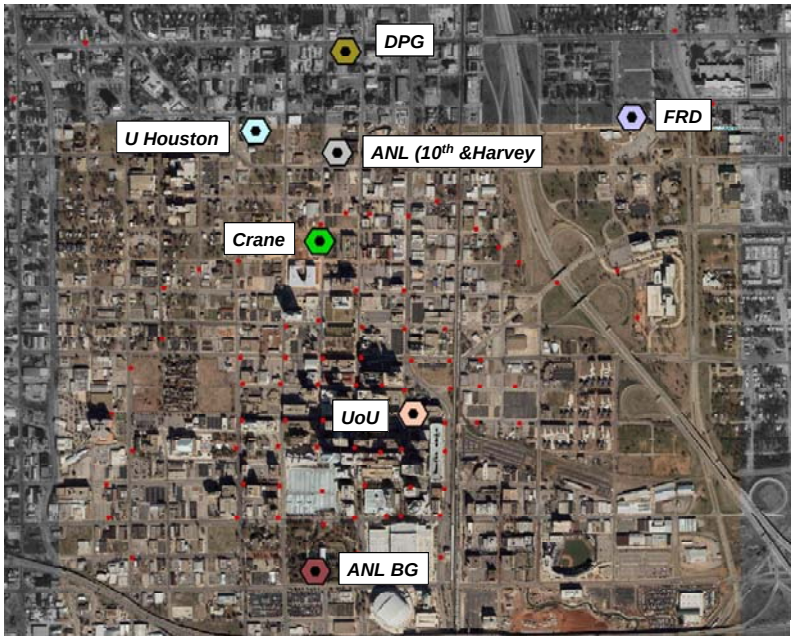


Figure 2-23. Close-In (CBD) Locations of Some Vertical Wind Profile Measurements During JU03 Using SODARs, mini-SODARs, and Sensors on a Crane

Red circles correspond to surface sampler locations.

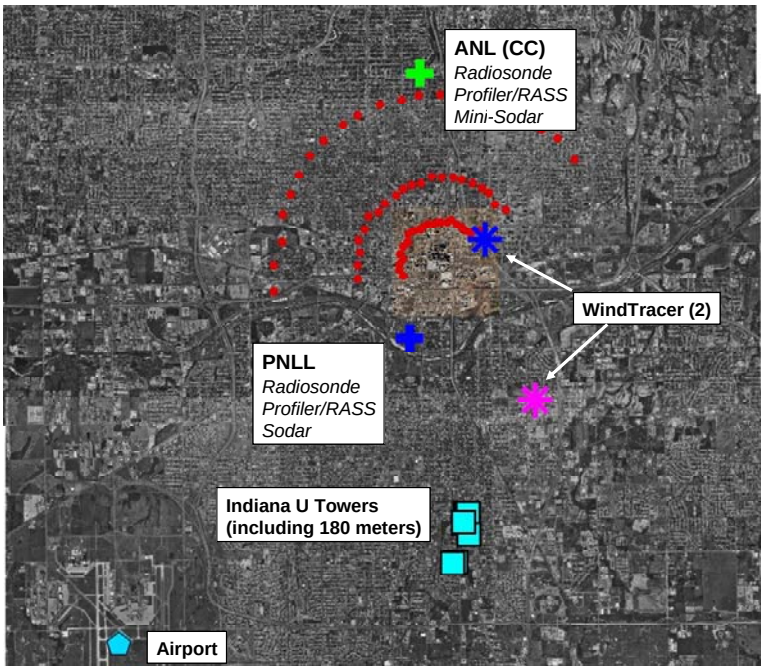


Figure 2-24. Locations of Upwind–PNNL and Indiana U Towers and Downwind–ANL(CC) Vertical Wind Profile Measurements, WindTracer Measurements and the Oklahoma City Airport

Red circles correspond to surface sampler locations.

The PNA MET option corresponded to combining both the SODAR and profiler (i.e., wind velocity as a function of height AGL) observations that were available at the upwind PNNL site (Figure 2-24).[5] Two previous studies of Urban HPAC [Refs. 2-4 and 2-5] suggested that a single measured upwind vertical profile can represent a satisfactory input to create reasonable urban predictions. Therefore, the PNA MET option allows for the further testing of this hypothesis. For the PNA MET option, the following should be noted: (1) profiler observations did not exist for the two IOP 1 releases and so only SODAR data are used for those PNA predictions and (2) attempts to create predictions of some releases with the PNA MET option caused HPAC to abort, citing a SWIFT error,[6] and therefore, for consistency of comparison, all PNA predictions used MC-SCIPUFF to create gridded wind fields as opposed to SWIFT. This is an important difference that will be discussed in the results section.

Figure 2-25 shows an example comparison of an Urban HPAC prediction to JU03, IOP 9, Release 1 observation using the PNA MET option and DM. Figure 2-26 shows another PNA-based set of comparisons, this time for all five Urban HPAC modes and IOP 5, Release 3.

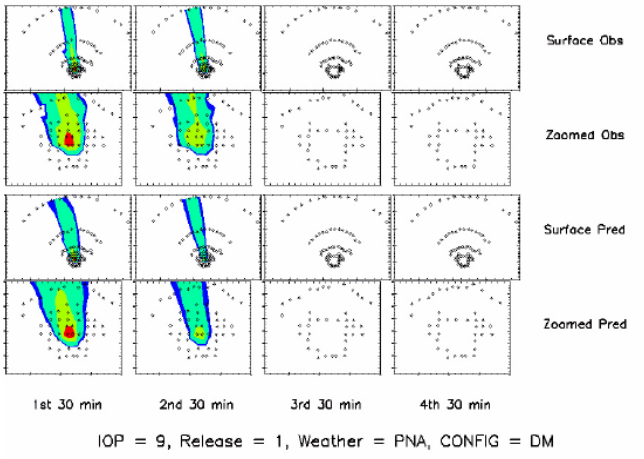


Figure 2-25. Comparisons of 30-Minute Average Concentration Contours for IOP 9, Release 1: Predictions Based on the PNA MET Option and the DM Urban HPAC – "PNA_DM"

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

---

5 PNA is short for "PNNL All," where "All" implies both the SODAR and profiler observations were used when available. The SODAR and profiler observations were combined using the meteorological editor resident within HPAC.

6 We were unable to discover the detailed cause of this SWIFT error but have passed along the information associated with these failures to SAIC, the contractor responsible for SWIFT integration in HPAC.

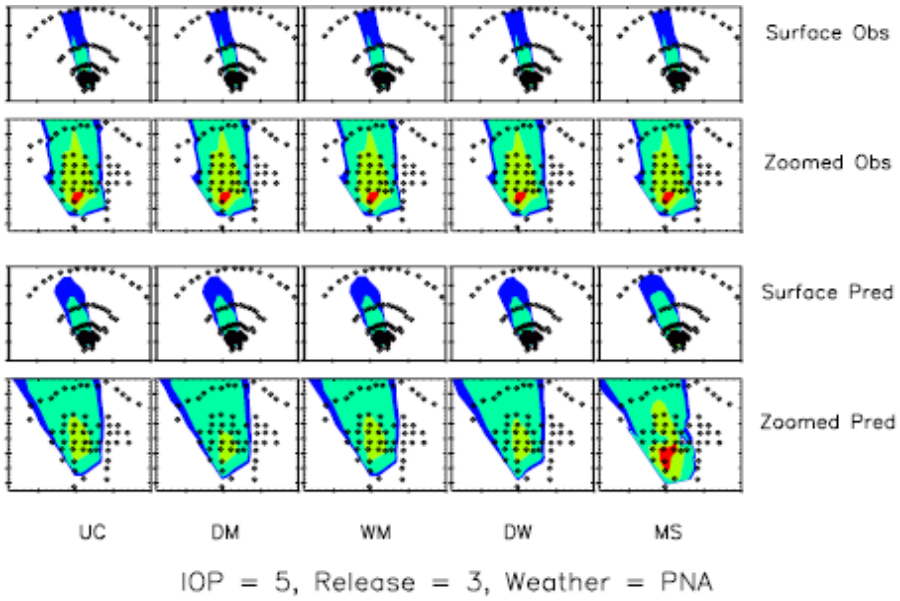


Figure 2-26. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 5, Release 3: All Predictions Based on the PNA MET Option

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

d. ACA

The ACA MET option corresponded to combining both the SODAR and profiler observations that were available at the downwind ANL (CC) site (Figure 2-24).[7] This MET option was considered particularly useful for comparison and contrast with PNA – the single upwind site option. For the ACA MET option, the following should be noted: (1) SODAR observations did not exist for the three IOP 3 releases and so only profiler data are used for those PNA predictions and (2) attempts to create predictions of some releases with the ACA MET option caused HPAC to abort, citing a SWIFT error, and therefore, for consistency of comparison, all ACA predictions used MC-SCIPUFF to create gridded wind fields as opposed to SWIFT. This is an important difference that will be discussed in the results section.

Figure 2-27 shows an example comparison of an Urban HPAC prediction to JU03, IOP 2, Release 2 observation using the ACA MET option and UC. Figure 2-28 shows another ACA-based set of comparisons, this time for all five Urban HPAC modes and IOP 4, Release 1.

---

7 ACA is short for "ANL (CC) All," where "All" implies both the SODAR and profiler observations were used when available. The SODAR and profiler observations were combined using the meteorological editor resident within HPAC.

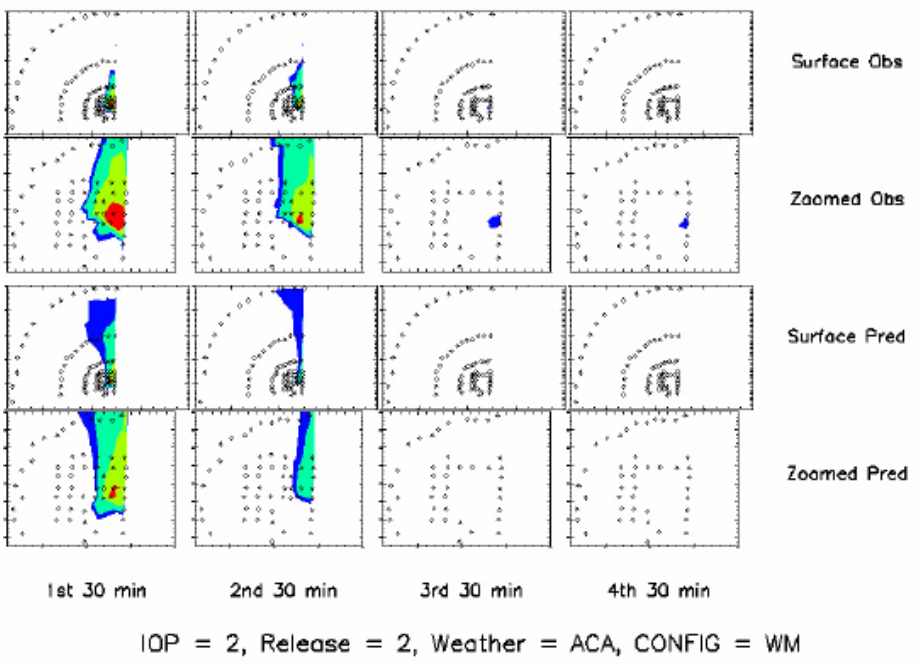


Figure 2-27. Comparisons of 30-Minute Average Concentration Contours for IOP 2, Release 2: Predictions Based on the ACA MET Option and the WM Urban HPAC Mode – "ACA_WM"

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

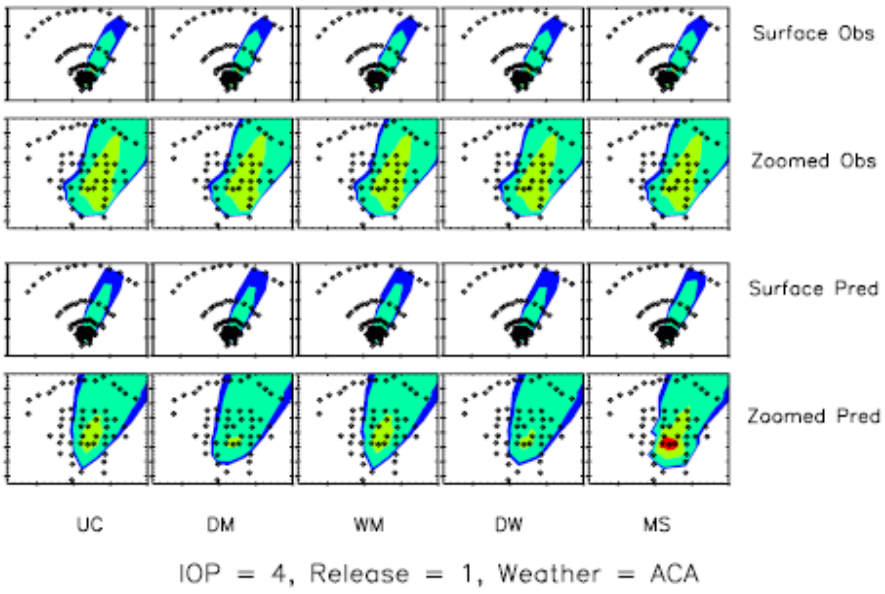


Figure 2-28. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 4, Release 1: All Predictions Based on the ACA MET Option

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

e. PO7

The PO7 MET option corresponded to a set of observations from a single location 40 meters AGL on the roof of the (upwind) Oklahoma City Post Office building (Figure 2-20). This option corresponds to using a single downtown, somewhat upwind, observation as input for the Urban HPAC predictions. During a previous study of urban atmospheric transport and dispersion ("Urban 2000" in Salt Lake City), it was found that using a single downtown building top measurement resulted in relatively degraded predictions in terms of fits to the observations when compared with the other MET options that were examined [Ref. 2-4]. Therefore, the PO7 MET option allows for the reconsideration of this concept, i.e., using a set of observations from a single building top as input for hazardous material transport and dispersion predictions, albeit this time a somewhat upwind building as opposed to a downtown building.

A PWIDS was located on the roof of the post office. This instrumentation collected wind speed and direction data every 10 seconds. As part of a preliminary study, we explored different approaches to processing this 10-second data for use as an Urban HPAC input. We created complete sets of predictions using the post office-based observations processed in the following ways:

1. Using the 10-second information directly as input
2. Scalar[8] averaging of 5 minutes of data and identifying the time as the midpoint
3. Vector averaging of 5 minutes of data and identifying the time as the midpoint
4. For each 5-minute interval, scalar averaging of the last 2 minutes of data and identifying the time as the endpoint
5. For each 5-minute interval, vector averaging of the last 2 minutes of data and identifying the time as the endpoint
6. Scalar averaging of 15 minutes of data and identifying the time as the midpoint
7. Vector averaging of 15 minutes of data and identifying the time as the midpoint
8. For each 15-minute interval, scalar averaging of the last 2 minutes of data and identifying the time as the endpoint

---

[8] Scalar and vector averaging refer to the manner in which wind speed and direction data are averaged. Scalar averaging corresponds to separately averaging wind speeds and wind directions over the time period of interest. Vector averaging corresponds to averaging the wind vectors over the time period of interest.

9. For each 15-minute interval, vector averaging of the last 2 minutes of data and identifying the time as the endpoint.

Some of these processing methods (4, 5, 8, and 9) were particularly consistent with somewhat standard meteorological practices. In general, these processing techniques resulted in predictions that were on average very similar.[9] The single exception to this rule was the first technique – "using the 10-second information directly as input." This option appeared to lead to marginally worse predictions (in terms of fits to the observations) than the other techniques. For the comparisons of this study, we chose the 7th technique described above – "vector averaging of 15 minutes of data and identifying the time as the midpoint." Hence, the shorthand notation is "PO7."

Figure 2-29 shows an example comparison of an Urban HPAC prediction to JU03, IOP 8, Release 3 observation using the PO7 MET option. Figure 2-30 shows another PO7-based set of comparisons, this time for all five Urban HPAC modes and IOP 10, Release 3.

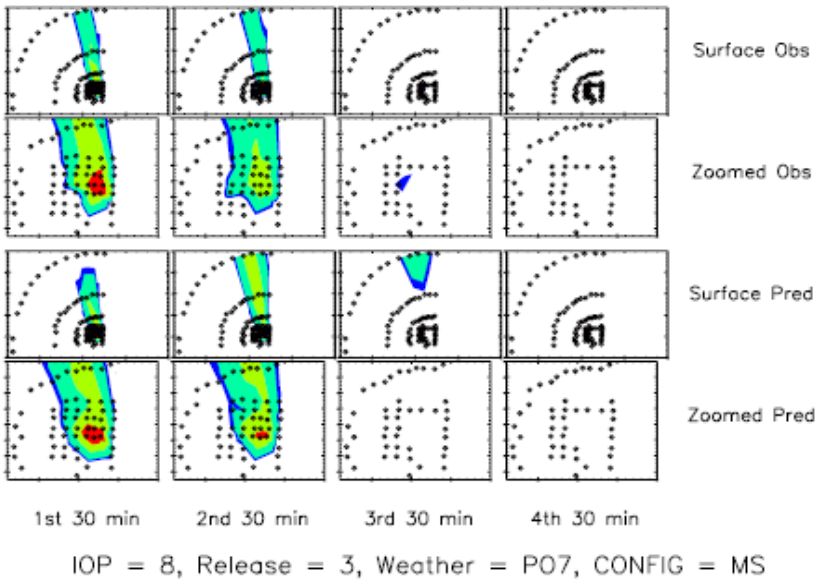


Figure 2-29. Comparisons of 30-Minute Average Concentration Contours for IOP 8, Release 3: Predictions Based on the PO7 MET Option and the MS Urban HPAC Mode – "PO7_MS"

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

---

9 We compared these sets of predictions using several metrics including a transport and dispersion measure of effectiveness described in the next section of this chapter and several standard statistics (also described in the next section).

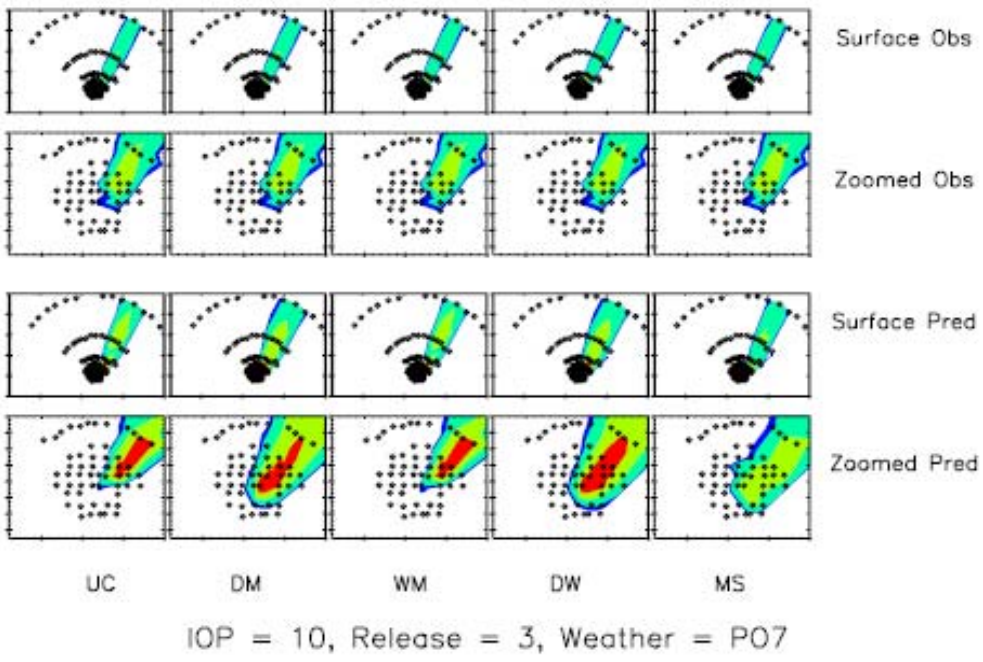


Figure 2-30. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 10, Release 3: All Predictions Based on the PO7 MET Option

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

f. Summary of *Compared* Sets of JU03 Predictions

Twenty-five sets of Urban HPAC predictions were generated as described in Table 2-4, which defines the shorthand notation that we used to describe each of the model runs. The HPAC "Export Project Folder" feature was used to generate electronic folders for each prediction. For UC, WM, DM, and DW, predictions were run for each IOP – that is, all releases (2 or 3) associated with that IOP. This results in 200 sets of predictions (4 urban modes × 5 MET options × 10 IOPs). In addition, for the MS Urban HPAC mode, individual files were created for each of the 29 releases. Therefore, an additional 145 sets of predictions (1 mode × 5 MET options × 29 releases) for the MS mode were created. These 345 folders contain all of the information needed, and in a convenient form, to allow anyone with the appropriate HPAC software to regenerate our predictions. For example, future sensitivity studies could be conveniently started using these predictions as a starting point.

In addition to the predictions described in Table 2-4, several other sets of MET options were investigated as described in the next subsection (A.2.g). In total, over 3,000 sets of predictions were created for this study.

Table 2-4. Shorthand Notations for the 25 Model Combinations (5 Transport and Dispersion Modes × 5 Meteorological Input Options) That Were Examined

	Transport and Dispersion Model Mode				
MET Options	**UC**	**WM**	**DM**	**DW**	**MS**
BAS	BAS_UC	BAS_WM	BAS_DM	BAS_DW	BAS_MS
BRB	BRB_UC	BRB_WM	BRB_DM	BRB_DW	BRB_MS
PNA	PNA_UC	PNA_WM	PNA_DM	PNA_DW	PNA_MS
ACA	ACA_UC	ACA_WM	ACA_DM	ACA_DW	ACA_MS
PO7	PO7_UC	PO7_WM	PO7_DM	PO7_DW	PO7_MS

g. Summary of Other MET Option Investigations

Several preliminary investigations of other MET options were begun as a part of this study. This section briefly describes some of these preliminary studies. Future reports will describe the details and results of these studies. This analyses and discussion of this subsection do not represent completed research. Rather, this section describes ongoing investigations and should provide a context for understanding the extent of the preliminary examinations that were a necessary part of this study.

i. Use of the ANL Botanical Gardens mini-SODAR ("BGS")

During preliminary studies of possible MET options, the downtown BGS site was explored. This site included a mini-SODAR that collected vertical profile information up to about 200 meters AGL. After creating predictions with the BGS MET option for all 29 releases, it was noted via a variety of metrics that the fit of predictions to observations was considerably worse than was the case for other MET options. Figure 2-31 compares predicted and observed "plumes" for an example daytime release using the BAS and BGS MET options. The prediction that used the BGS option appears to travel in a direction inconsistent with the observed (and BAS-predicted) plume. Figure 2-32 shows a portion of a table that lists the actual observations associated with the BGS air profile observations for IOP 6, Release 2. This table confirms the lower-altitude wind direction, from the south-southeast, that is reflected in the BGS prediction but not the observations.

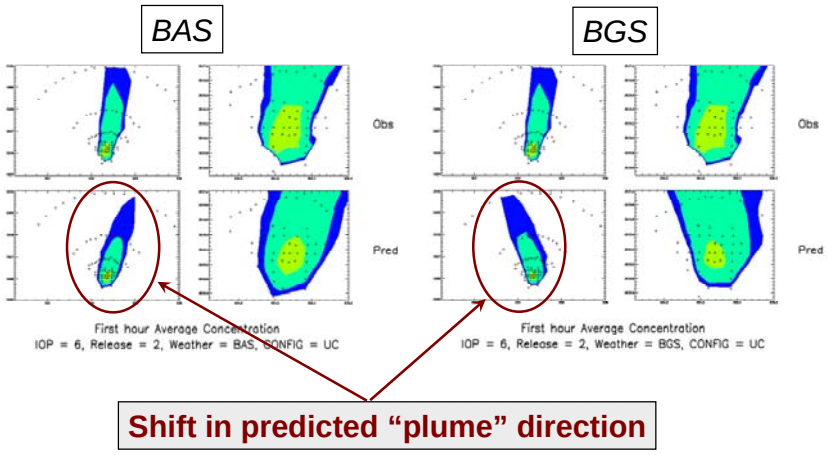


Figure 2-31. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 6, Release 2 (Daytime): Predictions Based on the BAS MET Option–Left and the BGS MET Option–Right

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

	Time = 16:00		Time = 16:15		Time = 16:30		Time = 16:45		Time = 17:00	
Height	WIDIR	WSPEED	WIDIR	WSPEED	WIDIR	WSPEED	WIDIR	WSPEED	WIDIR	WSPEED
5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999
10	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999
15	161	2.3	155	2.3	166	1.9	161	2.4	168	2.2
20	170	2.7	159	2.9	181	2.7	164	3.7	178	2.8
25	183	2.9	168	3.1	190	2.8	172	3.6	180	3.4
30	184	3.2	178	3.2	185	3.4	177	4.1	189	3.7
35	184	3.6	177	3.6	180	3.7	177	4.3	189	4.2
40	182	4.1	177	4.2	177	4.4	175	5.1	186	5
45	183	4.6	178	4.3	176	4.9	175	5.5	188	5
50	183	4.9	184	4.6	174	4.8	173	5.9	193	5.2
55	190	4.6	177	5.1	181	5.5	170	5.7	193	5.4
60	184	5.3	177	4.9	176	5.7	172	6	194	5.7
65	181	5.1	169	4.6	177	5.5	170	6.2	195	5.5
70	182	5.4	170	5.1	175	6.2	169	6.4	195	5.8
75	183	5.4	168	5.2	169	6.5	173	6.1	197	6.1
80	172	6.1	162	5.9	169	6.7	174	6.2	195	6
85	170	6.5	164	5.8	168	6.7	171	6.4	200	6
90	174	5.8	158	5.6	168	6.8	170	6	200	5.6
95	173	6	159	5	172	6.4	170	7.2	199	6.4
100	159	7.5	152	6.4	174	7.6	180	7	194	6.9

Winds show slight shift towards South-Southeast relative to BAS and observations. Wind speeds mostly monotonically increase with altitude

Figure 2-32. Relevant Excerpts from BGS MET Option Mini-SODAR Files for IOP 6, Release 2 (Daytime)

Figure 2-33 compares predicted and observed "plumes" for an example nighttime release using the BAS and BGS MET options, and Figure 2-34 shows a portion of a table that lists the actual observations associated with the BGS air profile observations for IOP 7, Release 1. Again, the lower-altitude BGS observations, this time very low wind speed, result in predictions that are inconsistent with the observed concentrations.

We hypothesize that the BGS in-city observations near the ground may be perturbed by the local urban environment and are not necessarily representative of the larger low-altitude wind flow. Further, the Urban HPAC model must accept these observations as representative and hence fit its near ground flow field (e.g., after SWIFT processing) to these non-representative observations. Therefore, we speculate that such low-altitude variable and non-representative measurements might be part of the cause of the poor relative performance of BGS-created predictions.

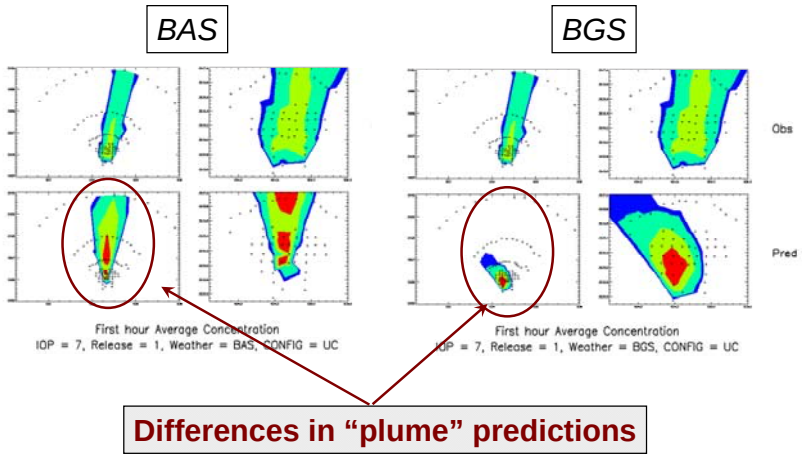


Figure 2-33. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 7, Release 1 (Nighttime): Predictions Based on the BAS MET Option–Left and the BGS MET Option–Right

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

	Time = 4:00		Time = 4:00		Time = 4:00		Time = 4:00		Time = 4:00	
Height	WDIR	WSPEED	WDIR	WSPEED	WDIR	WSPEED	WDIR	WSPEED	WDIR	WSPEED
5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999
10	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999
15	152	1.8	132	1.3	104	1.4	97	1.4	100	1.6
20	164	1.4	129	0.4	120	0.8	131	1.2	124	1.6
25	146	0.7	137	0.9	128	1.6	135	1.6	148	1.6
30	162	1.2	178	0.9	164	1	165	1	153	1.3
35	157	1	150	1.1	149	0.8	145	0.7	136	1
40	168	2	159	1.6	154	1.7	151	1.6	153	1.8
45	172	2.1	178	2.2	181	2.1	168	2.2	165	2
50	178	2.3	186	1.9	188	2.3	181	2.4	174	2.1
55	180	2.9	194	2.3	186	3.1	187	2.7	180	2.9
60	178	3.6	185	3	188	3.9	185	3.5	181	3.1
65	187	3.4	188	2.8	188	4.3	186	3.8	180	3.6
70	190	4.3	196	3.5	188	5.4	190	4.3	190	3.6
75	185	5.6	197	4.8	191	5.7	190	5	193	4.4
80	181	6	190	5.9	193	6	190	5.7	197	4.9
85	184	6.9	192	6	195	6.1	188	6.1	194	5.5
90	181	7.2	193	6.2	192	6.7	186	6.7	197	6.3
95	183	7.7	193	5.7	190	7.3	192	6.9	195	6.8
100	183	7.8	193	7	193	7.3	193	7.6	193	7.2
105	183	8	194	6.9	193	7.7	193	7.6	193	7.4

Low altitude winds show slight shift towards North-North-West. Wind speeds reach a minimum around 20-30 AGL

Figure 2-34. Relevant Excerpts from BGS MET Option Mini-SODAR Files for IOP 7, Release 1 (Nighttime)

ii. Exclusion of Low-Altitude Vertical Profile Observations

In part to explore the above hypothesis, we further examined the predictions that could be created from a few vertical measurements by cutting off the observations below threshold altitudes. For example, sets of JU03 predictions were created using the PNNL SODAR (Figure 2-24) but eliminating meteorological observations below 50 meters (set 1), 70 meters (set 2), 100 meters (set 3), 150 meters (set 4), 250 meters (set 5) and 350 meters (set 6). In addition, sets of predictions were created using the Dugway SODAR (Figure 2-22) but eliminating observations below 15 and 30 meters. Preliminary examination suggests that at night, removing the SODAR measurements below ~70-100 meters leads to improved Urban HPAC predictions of JU03. Additional hypotheses have resulted from this preliminary effort:

- Is there (typically) a substantially different flow at lower altitude versus higher altitude in Oklahoma City at night?
- Is the flow in the city "separated" from outside flow?

Additional research is required to answer these questions. Figures 2-35 and 2-36 show, as a function of time (at night) and altitude, observed wind vectors from the BGS mini-SODAR (near the city center) and the PNNL SODAR (about 1.6 km upwind), respectively. Significant differences in the low altitude and high altitude wind directions are seen and the variability of the lower altitude measurements is clearly seen. Future reports will document the results of this ongoing research.[10]

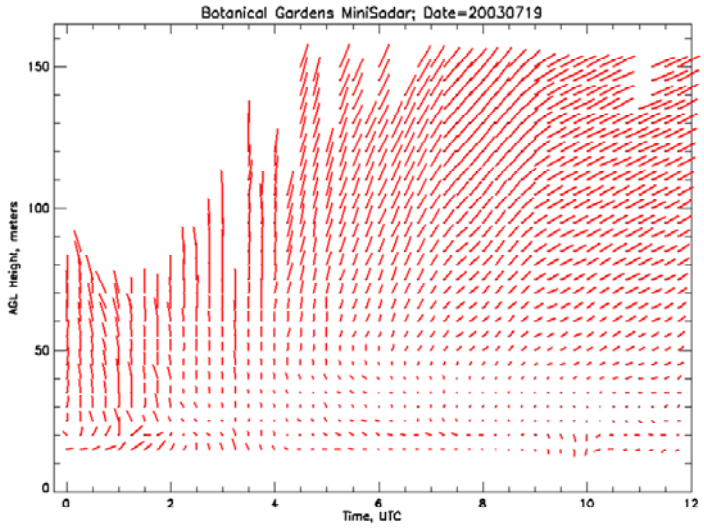


Figure 2-35. BGS Wind Vector Profiles as a Function of Time (at night) and Altitude

[10] Initial results of this study are reported in Ref. 2-6.

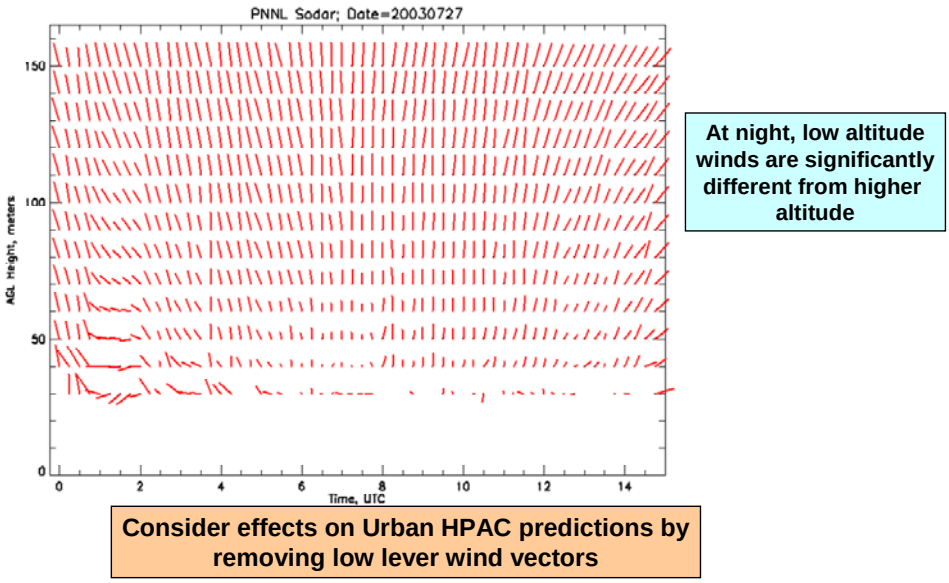


Figure 2-36. PNNL Wind Vector Profiles as a Function of Time (at night) and Altitude

iii. Use of LiDAR-Based Observations

Some have hypothesized that better spatially and temporally resolved meteorology could lead to improved predictions of hazards. Intuitively, if one knew the meteorological conditions precisely at any point in time and space, then one could significantly reduce the areas of false positive and false negative.[11] In recent years, processed LiDAR-based meteorological observations have been suggested as a means to improve model predictions by providing precise, highly resolved observations of the local wind field.

During JU03, two LiDARs were used to collect meteorological information. Two groups – NCAR and the Army Research Laboratory – participated in JU03 and collected LiDAR-based observations and could process these observations into HPAC-ready gridded files. IDA contacted both organizations and NCAR provided LiDAR-based meteorological observations for IOP 2. NCAR also provided real-time four dimensional data assimilation (RTFDDA) numerical weather information for IOPs 2, 3, 4, 5, 6, 9, and 10. Using their Variational LiDAR Assimilation System (VLAS), which assimilates high-resolution Doppler LiDAR data and other meteorological data, NCAR created observations for 13 "virtual towers" for IOP 2 and a full gridded field with a spatial resolution of 100 meters.

---

[11] Section B of this chapter describes a user-oriented measure of effectiveness and discusses the notion of false positives and false negatives in the context of T&D model evaluation.

Figure 2-37 shows the locations of the Army Research Laboratory LiDAR ("ARL LIDAR") and the Arizona State University LiDAR ("ASU LIDAR used by NCAR"). The locations of the 13 virtual towers created via VLAS-processing of LiDAR data by NCAR are also shown in Figure 2-37. For IOP 2, Urban HPAC predictions were created using the following three MET options: (1) RTFDDA, (2) 13 VLAS-generated virtual towers, and (3) full gridded VLAS-based wind vectors.[12]

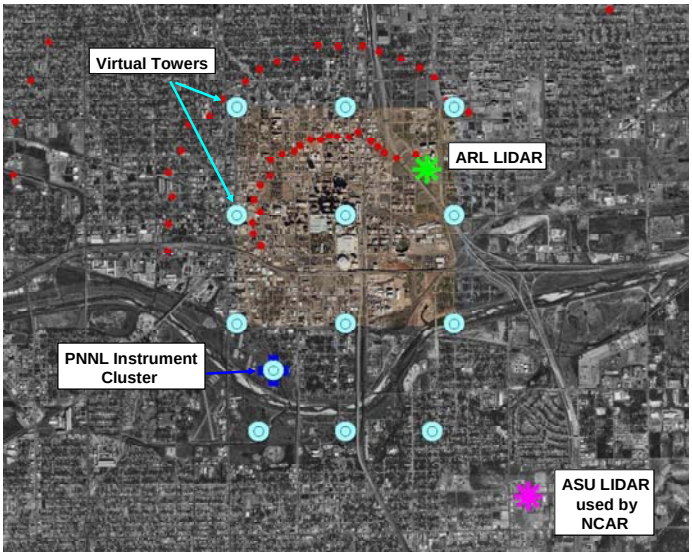


Figure 2-37. Locations of JU03 LiDAR Instrumentation, PNNL Cluster (e.g., SODAR and Profiler), and 13 Virtual Towers Created Via VLAS Processing of LiDAR Measurements

Figure 2-38 shows wind vectors as a function of location for the three types of meteorological input that was considered for this preliminary study: (1) RTFDDA, (2) VLAS-based virtual towers, and (3) full VLAS-based grid. The differences between wind fields are readily noticeable. The RTFDDA winds are mostly uniform, the virtual tower winds are relatively variable (maybe even erratic), and the full VLAS grid shows relatively detailed structure. Comparisons of predicted and observed concentrations are shown for the RTFDDA MET option in Figure 2-39 for IOP 2, Release 2. The overall plume direction appears reasonable, although much of the actual plume went off the sampler grid for this particular release. Figures 2-40 and 2-41 show the comparative results for IOP 2, release 2 predictions based on the VLAS-base virtual towers and full VLAS MET options, respectively. At least for this release, the predictions based on

[12] For these predictions, "terrain" and "landcover" were turned on within Urban HPAC and SWIFT was invoked.

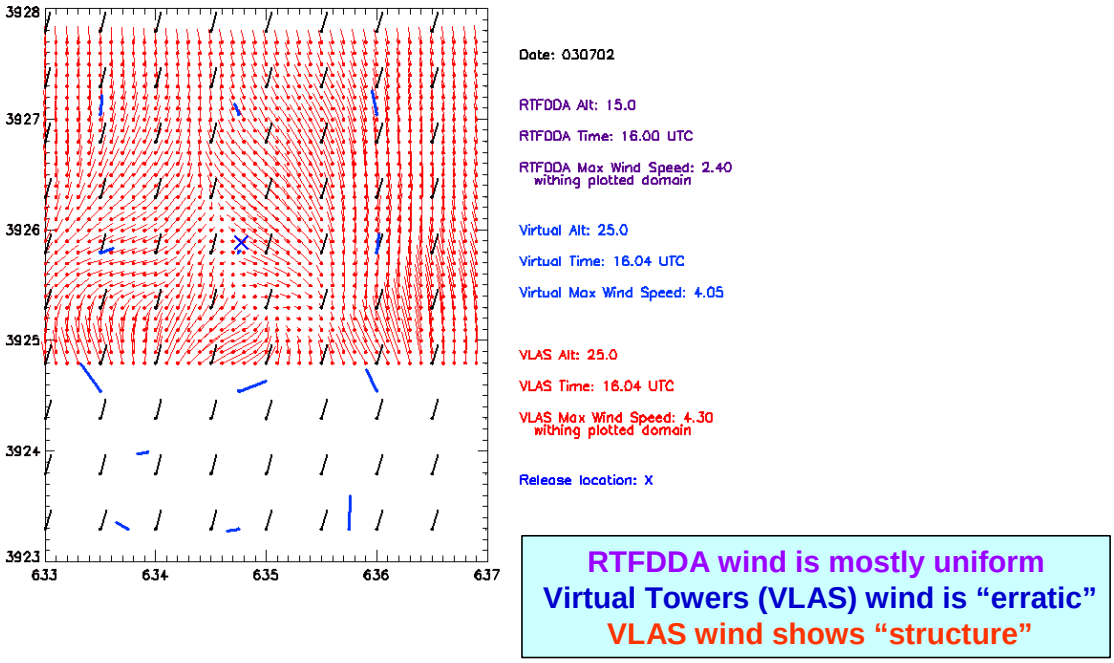


Figure 2-38. Comparison of Wind Vectors as a Function of Location: RTFDDA at 15 meters, VLAS-Based Virtual Towers at 25 meters, and Full VLAS-Based Grid at 25 meters

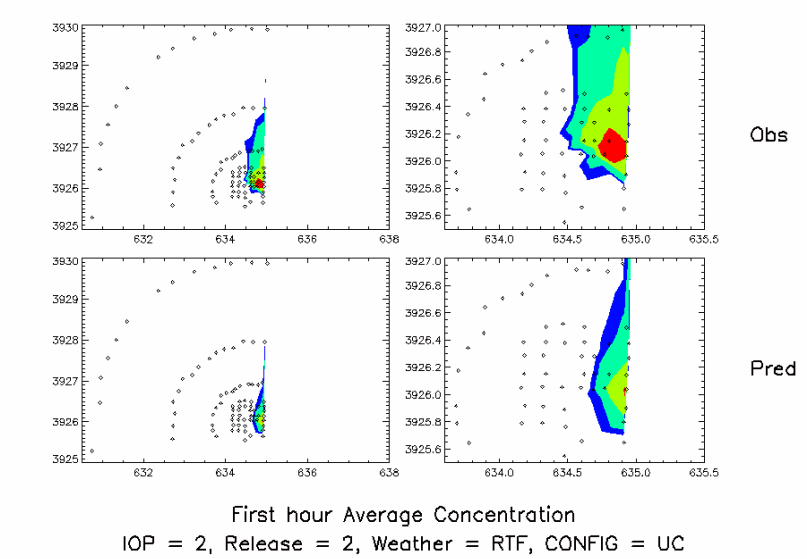


Figure 2-39. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 2, Release 2: Predictions Based on the RTFDDA MET Option and the UC Urban HPAC Mode

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

VLAS-processed LiDAR result in comparatively poorer predictions than the RTFDDA based predictions. This appears to be a result of the detailed ("erratic") structure

associated with the LiDAR based input. We plan to further research this unexpected finding and report results in future studies (e.g., could inherent mismatches in the averaging times associated with VLAS-processed measurements and the expected Urban HPAC input be a partial cause of this result?)

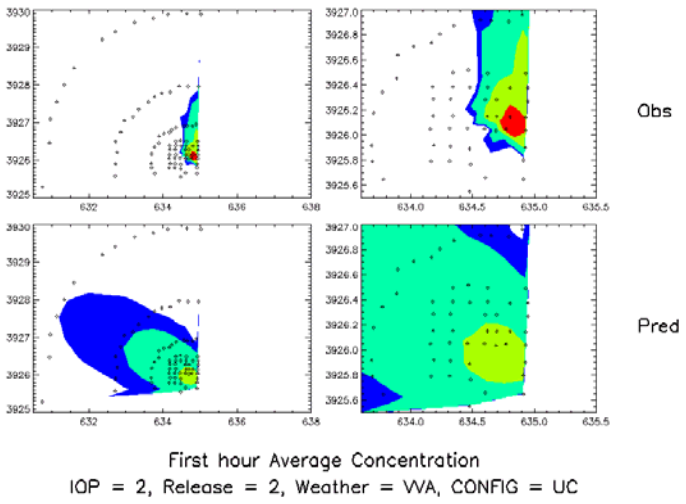


Figure 2-40. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 2, Release 2: Predictions Based on the 13 VLAS-Based Virtual Towers MET Option and the UC Urban HPAC Mode

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

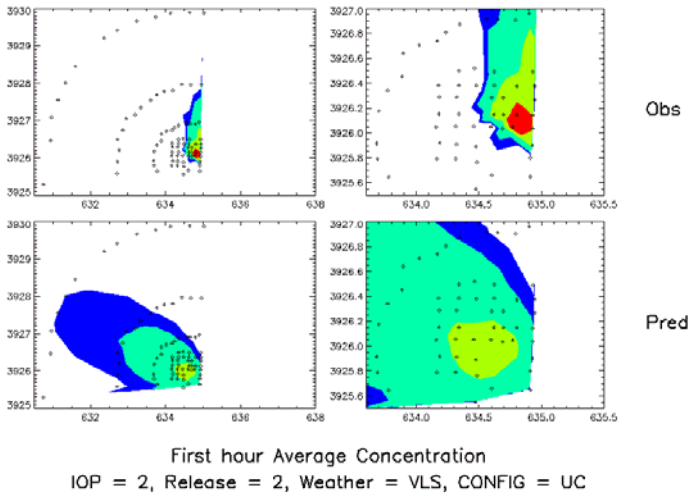


Figure 2-41. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 2, Release 2: Predictions Based on the Full VLAS MET Option and the UC Urban HPAC Mode

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

iv. Use of CCAM and MEDOC-Based Files

The Australian Defense Science and Technology Office (DSTO) made us aware of a regional climate model developed by Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO) known as the Conformal Cubic Atmospheric Model or CCAM. At the request of DSTO scientists and with the approval of our DTRA/Joint Science and Technology Office (JSTO) sponsor, we created CCAM-based predictions of JU03 using Urban HPAC. DSTO scientists provided HPAC-ready CCAM forecasts of JU03 to IDA at three spatial resolutions – 1 km ("C01"), 8 km ("C08"), and 60 km ("C60") – resulting from the original "nested" grids. The CCAM MET option comes in a format (meteorological data format – "MEDOC") that already accounted for terrain and does not allow for the running of the SWIFT processor. In this case, the gridded wind fields are used directly by HPAC-SCIPUFF. For comparisons we also obtained GCAT MEDOC files from NGIC and created Urban HPAC predictions. GCAT MEDOC-based predictions differ from the previously discussed BRB-based predictions because SWIFT is run in the case of BRB (but not in the case of GCAT MEDOC).

Our preliminary analysis found that Urban HPAC predictions can differ substantially depending on the format of the input MET option, e.g., MEDOC versus profile-based formats. The reason for these differences is currently being investigated and will be reported in future studies. With respect to the GCAT and CCAM MEDOC-format MET options, our examinations, which included comparisons of several metrics and formal hypothesis testing, suggest the following [Ref. 2-7]:

- GCAT MEDOC ("GCM") appears to be a slight improvement over CCAM ("C01," "C08," and "C60"). Importantly, while CCAM corresponds to a numerical weather forecast, the GCAT option (as used here) represented a numerical assimilation – that is, local observations were used to modify the GCAT-computed wind field. Therefore, the slight improvement in GCAT over CCAM can be understood in terms of assimilation versus a true forecast.
- CCAM-based predictions do not seem to be particularly sensitive to different grid resolutions (C01 versus C08 versus C60).
- CCAM and GCAT MEDOC predictions behave substantially different from predictions that use SWIFT or MC-SCIPUFF to process the meteorological input information.
- CCAM and GCAT MEDOC winds appear to transport the plumes too quickly compared to plumes resulting from predictions that use SWIFT and MC-SCIPUFF processed meteorological information.

Figure 2-42 shows an example comparison of predictions and observations (IOP 3, Release 3) for CCAM-based, GCAT MEDOC-format-base (GCM), and BRB MET options. CCAM-based predictions appear most similar to GCM-based predictions. Figure 2-43 provides evidence that CCAM-based predictions appear to transport the predicted plume substantially faster than the observed plume.

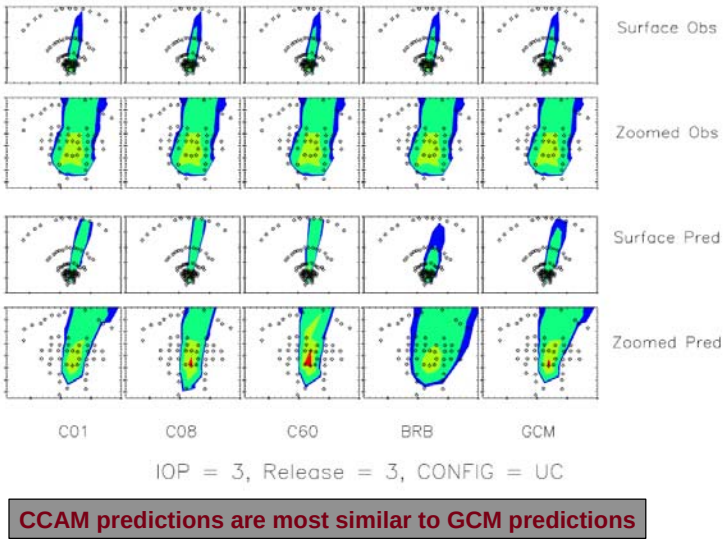


Figure 2-42. Comparisons of 1-Hour Average Concentration Contours for the First Hour for IOP 3, Release 3: All Predictions Based on the UC Urban HPAC Mode

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

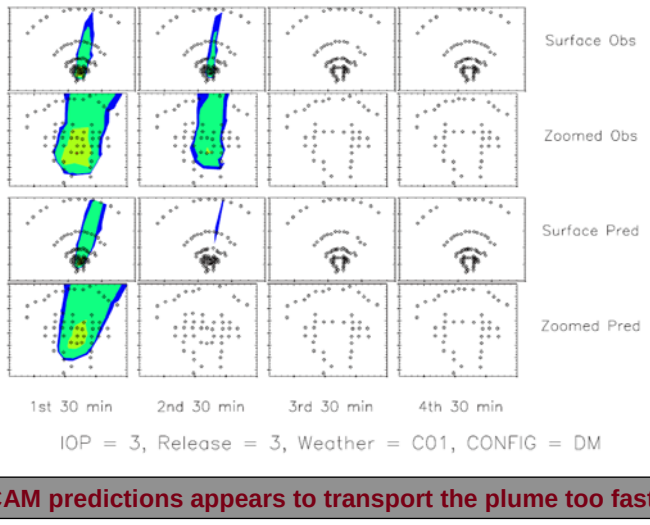


Figure 2-43. Comparisons of 30-Minute Average Concentration Contours for IOP 3, Release 3: Predictions Based on the CCAM 1 km Resolution (C01) MET Option and the DM Urban HPAC Mode

Contour levels are 25 (dark blue), 50 (light blue-green), 500 (green), and 500 ppt (red).

B. DESCRIPTION OF METRICS EXAMINED

In general, model validation efforts require specific measures of effectiveness to define metrics by which field trial observations and predictions can be compared. It is also helpful if model accreditation includes metrics that relate "operational" use of the model to field trial experiments. Such metrics give a certain degree of confidence to users with respect to how closely the model approximates the real world in their particular situation.

1. User-Oriented Measure of Effectiveness (MOE)

Previously, a user-oriented MOE [Ref. 2-8] has been developed and several applications have been described [Refs. 2-4 through 2-17]. A fundamental feature of any comparison of hazard prediction model output to observations is the over- and under-prediction regions. We define the false negative region where a hazard is observed but not predicted, and the false positive region where a hazard is predicted but not observed. Figure 2-44 shows one possible interpretation of these regions – the observed and predicted areas in which a prescribed dosage is exceeded. This view can be extended to consider the marginal over- and under-predicted values, as will be discussed below. In any case, numerical estimates of the false negative region (A_{FN}), the false positive region (A_{FP}), and the overlap region (A_{OV}) characterize this conceptual view. Although Figure 2-44 notionally illustrates physical areas to construct MOE components, the computation of the MOE does not necessarily require estimated areas and, hence, area interpolation. In fact, for this study, no interpolations were used to compute the MOE values.

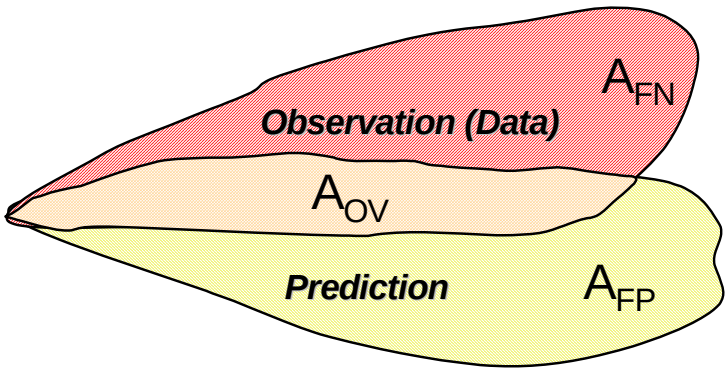


Figure 2-44. Conceptual View of Overlap (A_{OV}), False Negative (A_{FN}), and False Positive (A_{FP}) Regions that Are Used to Construct the User-Oriented MOE

The MOE that we consider has two dimensions. The x-axis corresponds to the ratio of the overlap region to the observed region and the y-axis corresponds to the ratio of the overlap region to the predicted region. When these mathematical definitions are

algebraically rearranged, one recognizes that the x-axis corresponds to 1 minus the false negative fraction and the y-axis corresponds to 1 minus the false positive fraction,

$$MOE = (x, y) = \left(\frac{A_{OV}}{A_{OB}}, \frac{A_{OV}}{A_{PR}} \right) = \left(\frac{A_{OB} - A_{FN}}{A_{OB}}, \frac{A_{PR} - A_{FP}}{A_{PR}} \right) = \left(1 - \frac{A_{FN}}{A_{OB}}, 1 - \frac{A_{FP}}{A_{PR}} \right) \quad (2\text{-}1)$$

where A_{FN} = region of false negative, A_{FP} = region of false positive, A_{OV} = region of overlap, A_{PR} = region of the prediction, and A_{OB} = region of the observation. This two-dimensional MOE includes directional effects; that is, the prediction of the location of a hazard and not just the shape and size of the plume is critical to obtaining a high MOE "score." From Eq. (2-1) it can be seen that MOE values along the "diagonal" of the two-dimensional MOE space (i.e., x = y) imply equal sizes (e.g., areas or amounts of material) of the prediction and the observation (i.e., $A_{PR} = A_{OB}$), even if the locations differ (Figure 2-45).

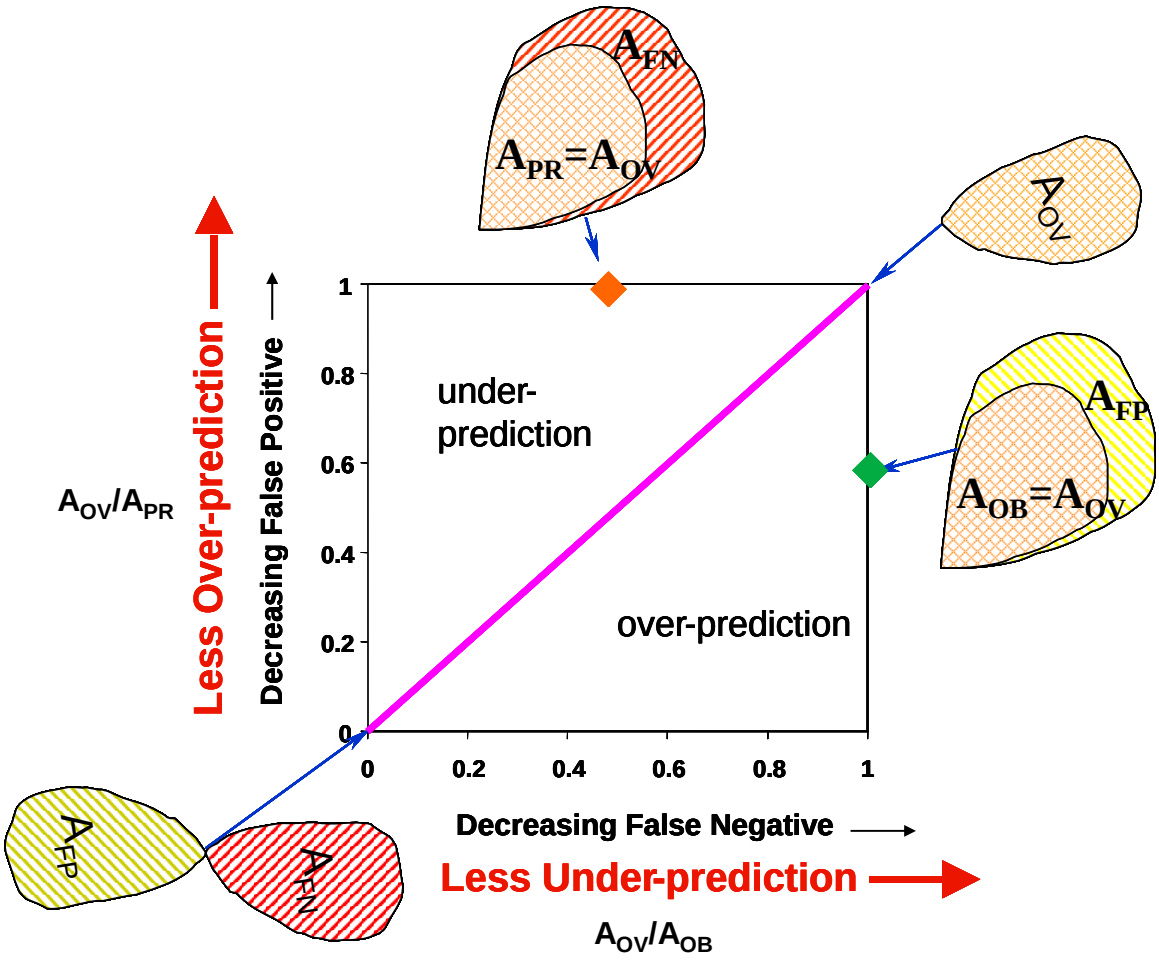


Figure 2-45. Key Characteristics of the Two-Dimensional MOE Space

A perfect model prediction leads to no false negative *and* no false positive, that is, complete and perfect overlap of the predictions and observations. Such a perfect model would have a two-dimensional MOE value of (1,1) as illustrated in Figure 2-45.[13] Additional discussion of the key characteristics of the two-dimensional MOE space can be found in Reference 2-8a.

The quantities A_{FN}, A_{FP}, and A_{OV} can be computed directly from the predictions and field trial observations paired in space and time. For the concentration-based MOE, the false positive region is the concentration predicted at the samplers but not observed. Therefore, for A_{FP} (as shown in Figure 2-46a), one first considers all of the samplers at which the prediction is of greater value than the observation. Next, one sums the differences between the predicted and observed average concentrations at those samplers (the yellow bars in Figure 2-46a). Based on the samplers that contained observed values that were larger than the predicted values, one can similarly compute A_{FN} (by summing the red bars in Figure 2-46a). A_{OV} is calculated by considering all samplers and summing the concentrations associated with the minimum predicted or observed value (the green bars in Figure 2-46a). These estimates can be made on a linear scale or on a logarithmic scale, as shown in Figure 2-46a.

In addition to applying the more general technique described above, one can compute an MOE value based on a prescribed threshold (concentration or dosage). First, one considers the predictions and observations at each of the samplers. If both the prediction and observation are above the threshold, it is considered overlap at that sampler (and the contributions to A_{OV}, A_{FN}, and A_{FP} from this sampler location are 1, 0, 0, respectively). If the prediction is below the threshold and the observation is above, a false negative is assessed at that sampler (and the contributions to A_{OV}, A_{FN}, and A_{FP} from this sampler location are 0, 1, 0, respectively). Similarly, a false positive is assessed when the prediction is above the threshold and the observation is not (and the contributions to A_{OV}, A_{FN}, and A_{FP} from this sampler location are 0, 0, 1, respectively). For the case of a specific sampler at which both the prediction and the observation are below the threshold, the values are assessed as 0, 0, 0 for the computation of the threshold-based MOE (consistent with the conceptual view illustrated in Figure 2-44). Figure 2-46b illustrates this procedure for an average concentration threshold of 25 ppt.

---

[13] A model prediction that completely misses the observation (for example, the "plume" goes in the exact opposite direction) would achieve an MOE value of (0,0).

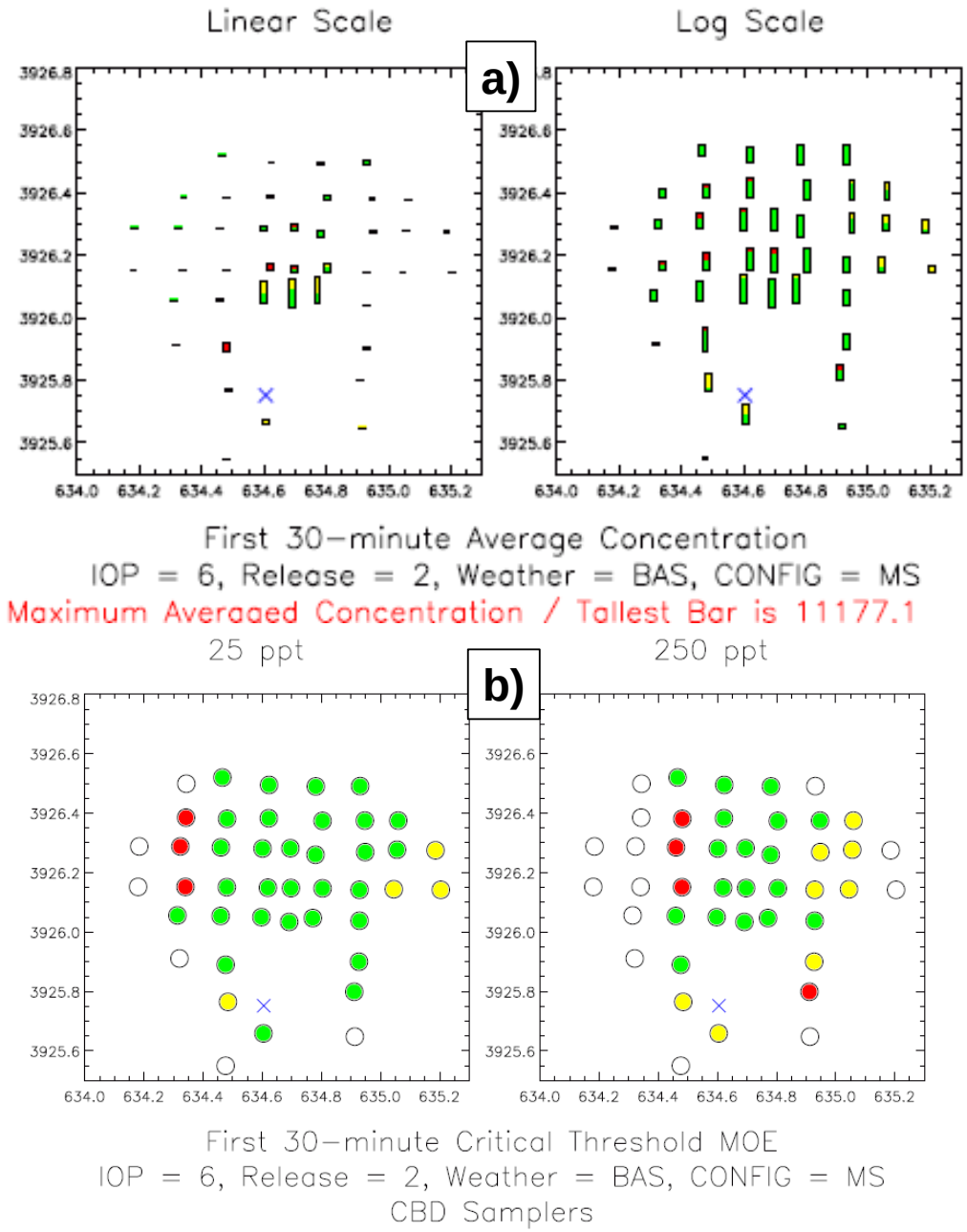


Figure 2-46. Illustration of MOE Component Computations for JU03 IOP 6, Release 2 for the CBD Sampler Locations With Predictions Based on the MS Mode Using the BAS Meteorological Input Option. a) Based on average concentrations – linear (left) and log (right) scales are shown. b) Based on a 25 ppt (left) and 250 ppt (right) average concentration thresholds.

The blue "X" corresponds to the release location, green circles indicate locations where both the observation and the prediction were above the threshold, clear (white) circles correspond to locations with both the observation and the prediction below the threshold, red circles indicate an observation above threshold and a prediction below threshold and finally, yellow circles correspond to sampling locations with predictions above threshold and observations below threshold. Thus, for the 250-ppt example shown above, A_{OV} = 18 samplers, A_{FN} = 4 samplers, and A_{FP} = 8 samplers.

In physical space (given interpolation of observations and predictions), this procedure approximately corresponds to assessing the MOE using a contour level (e.g., as illustrated conceptually in Figure 2-44).

We examined MOE values based on concentration thresholds of that exceeded 25, 250, and 2500 ppt. These three threshold values correspond to 5, 50, and 500 × the estimated background. The lower thresholds correspond to levels that are reasonably detectable *and* cover enough samplers to allow for a notion of the plume size, shape, and direction. This allows one to create MOE values, at least for the lower threshold values, that provide information on the model's ability to predict the "hazard" area. For some releases, the highest thresholds that were examined – 2500 ppt – produced situations where all samplers and predictions were below the threshold. For these cases, the MOE value was defined to be (1,1). Although examined, none of our discussed results are based on these "highest" thresholds. The above discussions imply that a total of four MOEs were examined – three based on thresholds and one based on summed concentrations.

2. Standard Statistics Computed

Several comparisons of transport and dispersion model predictions to field trial observations have demonstrated a variety of measures and issues [Refs. 2-18 through 2-25]. The statistical measures that were examined in this study include fractional bias (FB), geometric mean bias (MG), normalized mean square error (NMSE), bounded normalized mean square error (BNMSE), normalized absolute difference (NAD), geometric mean variance (VG), linear Pearson correlation coefficient (R), correlation coefficient based on logarithms (R_{ln}), root mean square error (RMSE), bias, fraction of predictions within a factor of 2 (FAC2), fraction of predictions within a factor of 5 (FAC5), and fraction of predictions within a factor of 10 (FAC10). These measures are defined below:

$$FB = \frac{\left(\overline{C_p} - \overline{C_o}\right)}{0.5\left(\overline{C_o} + \overline{C_p}\right)} \tag{2-2}$$

$$MG = \exp\left(\overline{\ln C_o} - \overline{\ln C_p}\right) = \exp\left(\overline{\ln\left(\frac{C_o}{C_p}\right)}\right) \tag{2-3}$$

$$NMSE = \frac{\overline{\left(C_p - C_o\right)^2}}{\overline{C_o}\;\overline{C_p}} \tag{2-4}$$

$$BNMSE = \frac{\sum_{i=1}^{n}\left(C_p^{(i)} - C_o^{(i)}\right)^2}{\sum_{i=1}^{n}\left(C_o^{(i)} + C_p^{(i)}\right)^2} \tag{2-5}$$

$$NAD = \frac{\sum_{i=1}^{n}\left|C_p^{(i)} - C_o^{(i)}\right|}{\sum_{i=1}^{n}\left(C_o^{(i)} + C_p^{(i)}\right)} \tag{2-6}$$

$$VG = \exp\overline{\left(\ln C_o - \ln C_p\right)^2} = \exp\left[\overline{\left(\ln\left(\frac{C_o}{C_p}\right)\right)^2}\right] \tag{2-7}$$

$$R = \frac{\overline{\left(C_o - \overline{C_o}\right)\left(C_p - \overline{C_p}\right)}}{\sigma_{C_p}\sigma_{C_o}} \tag{2-8}$$

$$R_{\ln} = \frac{\overline{\left(\ln C_o - \overline{\ln C_o}\right)\left(\ln C_p - \overline{\ln C_p}\right)}}{\sigma_{\ln C_p}\sigma_{\ln C_o}} \tag{2-9}$$

$$RMSE = \sqrt{\overline{\left(C_p - C_o\right)^2}} \tag{2-10}$$

$$Bias = \overline{C_p} - \overline{C_0} \tag{2-11}$$

$$\text{FACx = fraction of data for which } \frac{1}{x} \le \frac{C_p}{C_o} \le x \tag{2-12}$$

where C = observation/prediction of interest (e.g., 30-minute average concentration or 2-hour dosage), C_p corresponds to model predictions, C_o corresponds to observations, a bar above the quantity (e.g., $\overline{C}$) denotes the average (e.g., of the average concentrations), σ_C = standard deviation, n = number of data points used in the comparisons, $C_o^{(i)}$ refers to the ith observed (e.g., 30-minute average) concentration (dosage), and similarly, $C_p^{(i)}$ refers to the ith predicted concentration (dosage), "i" indexes the paired observations and predictions, and for this study x = 2, 5, and 10.

FB and MG measure the systematic bias in a model in terms of absolute differences and ratios, respectively. FB > 0 implies over-prediction and FB < 0 implies under-prediction. NMSE, BNMSE, NAD, and VG measure the scatter associated with

the predictions relative to observations (with VG measuring scatter associated with ratios). Perfect agreement with a set of observations would result in MG, VG, R, R_{ln}, FAC2, FAC5, and FAC10 = 1.0; and FB, NMSE, BNMSE, NAD, RMSE, and bias = 0.0.

For this study, the above measures were computed and examined and the inherent limitations of each were recognized. It is expected that, taken together, the measures described above can provide good insight into the relative performance of a given model. Although all of the above measures were computed and examined, for simplicity, most of the discussions of results in this paper will be based on FB (our chosen measure of bias), NAD and NMSE (our chosen measures of scatter), and the previously described MOE.[14] NMSE assesses scatter between observations and predictions but considers normalized *squared* differences as opposed to *absolute* differences. These differences in scatter metrics result in somewhat different sensitivities (previously discussed in Ref. 2-4).

In Chapter 3, we consider three metrics as measuring scatter between predictions and observations (at least partially) – NAD, NMSE, and the concentration-based MOE – and scatter is considered an important measure of model predictive performance. When applied to observations and predictions paired in space and time, these scatter-based metrics get at the question of how well the model predicted the location and timing (at least for 30-minute averages here) of the observations. For our Chapter 3 discussions, we demand that at least two of the three scatter-related metrics reveal statistically significant improvements before we assess a comparative result in terms of one mode outperforming another for a particular situation. In part, this allows us to assess the robustness of our conclusions that rely, at least partially, on relative differences between scatter metric values. It should be noted that for the three scatter-related metrics discussed above, we examined 600 hypothesis tests (600 = 3 metrics × 10 comparisons × 4 conditions × 5 MET options), and if, for example, one demanded a significance level of 0.05 before highlighting a difference as an improvement, one would expect about 30 false significance findings. However, by demanding that two of three differences result in statistical significance, these false positive findings ("Type I error") are greatly reduced.

---

[14] Approximate 0.99 confidence regions were computed for the MOE values using the bootstrap [Ref. 2-26] percentile method (10,000 bootstrap samples) and resampling MOE two-dimensional vectors [Refs. 2-4 and 2-8a].

C. METHODOLOGY FOR COMPARSIONS OF PREDICTIONS AND OBSERVATIONS

1. Scope of Comparisons

For the comparisons discussed here, we considered 120 surface samplers and 29 independent releases (i.e., 3,480 pairings near the surface). For each release, we examined the following six sets of samplers separately: all surface samplers, CBD (55 samplers), all arcs (65 samplers), 1 km arc only (23 samplers), 2 km arc only (21 samplers) and 4 km arc only (21 samplers). With respect to the averaging times associated with the 2-hour monitoring period after the start of the releases, we examined the eight 15-minute, four 30-minute (e.g., see Figure 2-43), two 1-hour, and single 2-hour time periods individually, and all 30-minute time periods together. We separately examined the daytime releases – IOPs 1-6, 17 releases – and the nighttime releases – IOPs 7-10, 12 releases (see Table 1-1). This yields 192 conditions (6 sampler sets × 16 time periods × 2 times of day) that were explored.

Combining the 4 MOEs and 13 standard statistics discussed previously with the 192 conditions discussed above leads to 3,264 measures that were computed for each Urban HPAC model prediction. As described in Table 2-4, 25 sets of model predictions (5 model modes × 5 MET options) were created implying a total number of computed metrics of 81,600. Although we have examined and retained all of these measures, our conclusions will be described in terms of only a few of these metrics.

Much of the analyses and discussion of Chapter 3 divides the sampler locations into two regimes – the CBD and the arcs. In a sense, examining the data and comparisons in this way serves as a surrogate for several potential considerations. For example, the examinations using the CBD observations focus on short-range, high-concentration performance while the analysis based on the arcs focus on the longer downwind distance and lower-concentration performance. It should be noted that the direct impact of some of the model components (e.g., UWM, UDM, and MSS) occurs at short range, with the SCIPUFF code computing the transport and dispersion outside of about (typically) 1 km, albeit using initial conditions "handed off" from the closer-in calculations. Finally, there is an expectation that the analyses of the more distant arc-based samplers might show more sensitivity to differences in wind direction accuracy relative to the CBD where the broad predicted plume might typically cover many of the close-in samplers.

2. Hypothesis Test Procedures: Computation of p-values

Given the very large set of comparisons suggested above, an objective way to logically identify differences between sets of model predictions was needed. We decided to compute p-values for the hypothesis testing of differences between estimated metrics associated with different model configurations as a semi-automated way to direct our search and allow us to comment on the relative performance of different sets of predictions.

For these data, it was important that for comparisons, metrics be "paired" by release, since the variance between releases was as large or larger as the variance between model configurations (which was associated with the comparisons of most interest). Significant differences between models were sought by comparing the statistic of interest for two models (at a time) for each of the 17 daytime or 12 nighttime releases. The statistic of interest, for example, FB, is paired by release (e.g., IOP 5, Release 2) for each prediction set.

For the one-dimensional metrics (e.g., the conventional statistics), the Permutation test with general scores [Ref. 2-27] – a nonparametric test similar to the Sign and Wilcoxon Signed-Rank tests [Ref. 2-28] – was used to compute p-values. For suitably low p-values, one can reject the null hypothesis of equivalence between the values being compared [Refs. 2-29 and 2-30]. The Permutation test uses the absolute magnitude of the paired differences and is generally more powerful than the Wilcoxon Signed-Rank test, which only uses the rankings of the paired data.

For equivalence between the metrics being compared, consistent with the null hypothesis of no significant difference, the signs would equally likely be positive or negative for each paired difference. For this test, one computes the permutation reference set (of size 2N, where N = the number of independent releases) using the actual differences as the general scores (as opposed to the counts for the Sign test or the sum of the ranks for the Wilcoxon-Signed Rank test). The 2N-sized reference set arises from the permutations of the sign (+ or -) that can be considered for each of the N differences. The positive or negative sum (from adding all of the differences associated with the positive or negative differences) then serves as the test statistic to be compared to the 2N reference set (based on all possible permutations of the signs) to generate a p-value that can be computed exactly or via a Monte Carlo procedure [Ref. 2-27]. For this study, approximate two-sided p-values (and associated 0.99 confidence intervals for the estimated p-values) based on 10,000 Monte Carlo runs are reported. The above

computations were accomplished within StatXact 4 [Ref. 2-27]. For these comparisons, we considered p-values ≤ 0.05 as significant.

For the two-dimensional MOE, the chosen hypothesis test procedure starts by computing vector differences between various paired (by release) model predictions. If two sets of model predictions were identical, then the 17 (or 12 for nighttime releases) vector differences would be (0,0). For this study, the null hypothesis is that the two models being compared are equivalent. Under this hypothesis, any MOE vector difference is expected to be equally likely to reside in any of the four quadrants, defined as positive x, positive y (++); positive x, negative y (+-); negative x, positive y (-+); and negative x, negative y (--). Given this null hypothesis, one tests how unlikely the observed result is by simulating the appropriate distributions in the following way. First, the quadrant with the most MOE vector differences is identified and the number of differences in that quadrant is noted – the maximum observed number for any quadrant. Next, results for equivalent model predictions are simulated by creating 100,000 samples of 17 drawn from the uniform integer distribution on {1, 2, 3, 4} – that is, a multinomial distribution with equal likelihood for each of 4 outcomes. These 100,000 samples of 17 correspond to our simulated vector differences for equivalent models. For each sample of 17, the numbers of "1s," "2s," "3s," and "4s" that were randomly selected are determined. The maximum observed number for any quadrant is then compared to the corresponding maximum number associated with each simulated sample. The number of simulated samples that contain a maximum value that is greater than or equal to the observed maximum is determined and denoted $N_{\geq}$. The estimated p-value is then computed as $N_{\geq}$ divided by 100,000. The 2-dimensional, 4-quadrant hypothesis test described here is a natural extension of the one-dimensional Sign test.

For sample size 17 (daytime releases), using the four-quadrant test described above for all possible combinations, the closest (but below) p-value to 0.05 is 0.0496, which occurs for a maximum of 9 vector differences in any one quadrant. For sample size 12 (nighttime releases), the closest (but below) p-value to 0.05 is 0.0112, which occurs for a maximum of 8 vector differences in any one quadrant. Therefore, when examining MOE differences, 9 vector differences in any one quadrant were used to indicate a statistically significant difference for the daytime release comparisons and 8 differences were used for the nighttime releases.

Clearly, the above hypothesis testing procedures (or some other simplifying approach) is necessary to aid the analyst in wading through the many possible comparisons. All such approaches potentially suffer from spurious rejections of the null

hypothesis (e.g., 5 percent of standard nominal 0.05 significance level tests will falsely reject a true null). Overall findings in this study are reinforced by consistency across different measures.[15] Furthermore, the benefit of these nonparametric hypothesis test procedures has been previously demonstrated [Ref. 2-4].

---

[15] The methods of Bonferroni, among others, can also be used to mitigate the problem of spurious rejections of the null hypothesis. The Bonferroni correction can be applied when there is concern that in doing more than one test in a particular study, the alpha level should be adjusted downward to consider chance. The alpha level is the chance of incorrectly declaring a difference to be true when it is actually due to chance. For example, in five hypothesis tests done with an alpha level of 0.05, the chance of finding at least one difference or relationship significant due to chance fluctuation equals 0.22. The Bonferroni method adjusts the alpha level downwards for each individual test to ensure that the overall risk given the number of tests remains at 0.05. Of course, in making this adjustment one accepts a higher risk that no difference is detected when in fact there is a difference. As mentioned above, this study relies on consistency across different metrics that often provide information on similar model performance features (e.g., for scatter, NAD, NMSE, and the concentration-based MOE are examined together). Reference 2-31 provides some critical comments on the application of the Bonferroni method.

REFERENCES

2-1. Capability Production Document for Joint Effects Model Increment, Director, Joint Requirements Office for CBRN Defense (JRO-CBRND), 27 June 2006.

2-2. Allwine, K. J., M. J. Leach, L. W. Stockham, J. S. Shinn, R. P. Hosker, J. F. Bowers, and J. C. Pace, 2004: Overview of joint urban 2003—An atmospheric dispersion study in Oklahoma City. *Symp. on Planning, Nowcasting and Forecasting in the Urban Zone*, Seattle, WA, Amer. Meteor. Soc., January 12-16.

2-3. Vandenberghe, F., R. Weingruber, M. Casado, S. Swartz, R. Sheu, M. Ge, A. Bourgeoi, T. Betancourt, S. Swerdlin, T. Warner. and R. Babarsky: The Global Climatology Analysis Tool, accessed web site on 12/15/06 13:58, http://128.117.192.211/projects/global/references/GCAT_description_short.htm

2-4. Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of transport and dispersion model predictions of the URBAN 2000 field experiment. *J. Appl. Meteor*., **43**, 829-846.

2-5. Warner, S., N. Platt, J. F. Heagy, J. E. Jordan, and G. Bieberbach, 2006: Comparisons of transport and dispersion model predictions of the mock urban setting test field experiment. *J. Appl. Meteor. and Climatology*, **45**, 1414-1428.

2-6. Platt N., J. T. Urban, S. Warner, and J. F. Heagy, 2006: Comparison of different meteorological options used for Urban HPAC predictions of Joint Urban 2003 field trials. *Tenth Annual George Mason University Transport and Dispersion Modeling Workshop,* Fairfax, VA, 18 pp, August 2006.

2-7. Platt N., S. Warner, J. T. Urban, and J. F. Heagy, 2007: Update on Urban HPAC Evaluations with joint Urban 2003 Field Trial Data. *TTCP-CBD Technical Panel 9 (TP9)) Annual Meeting,* Porton Down, UK, 60 pp, February 2007.

2-8. a) Warner, S., N. Platt, and J. F. Heagy, 2004: User-oriented two-dimensional measure of effectiveness for the evaluation of transport and dispersion models. *J. Appl. Meteor*., **43**: 58-73. b) Warner, S., N. Platt, and J. F. Heagy, 2001: User-oriented measures of effectiveness for the evaluation of transport and dispersion models. *Proceedings of the 7th Int'l Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, Belgirate, Italy, 28-21 May 2001, pages 24-29. c) Warner S., N. Platt, J. F. Heagy, S. Bradley, G. Bieberbach, G. Sugiyama, J. S. Nasstrom, K. T. Foster, and D. Larson, 2001: *User-Oriented Measures of Effectiveness for the Evaluation of Transport and Dispersion Models,* IDA Paper P-3554. (Available electronically at DTIC STINET ada387239.)

2-9. a) Warner, S., N. Platt, and J. F. Heagy, 2005: Comparisons of transport and dispersion model predictions of the European tracer experiment: area-based and population-based measures of effectiveness. *Atmospheric Environment*, **39**, 4425-4437. b) Warner, S., N. Platt, and J. F. Heagy, 2004: *Comparisons of Transport and Dispersion Model Predictions of the European Tracer Experiment: Area-Based and Population-Based Measures of Effectiveness,* IDA Paper P-3915. (Available electronically at DTIC STINET ada427807.)

2-10. a) Warner, S., N. Platt, and J. F. Heagy, 2004: Application of user-oriented MOE to transport and dispersion model predictions of the European tracer experiment. *Atmospheric Environment*, **38**, 6789-6801. b) Warner S., Platt, N., and Heagy, J. F., Warner, S., N. Platt, and J. F. Heagy, 2003: *Application of User-Oriented MOE to Transport and Dispersion Model Predictions of the European Tracer Experiment,* IDA Paper P-3829. (Available electronically at DTIC STINET ada419433.) c) Platt, N., S. Warner, and J. F. Heagy, 2004: Application of user-oriented MOE to transport and dispersion model predictions of ETEX. *Proceedings of the 9th Int'l Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, 1-4 June 2004, pages 120-125.

2-11. Heagy, J. F., N. Platt, and S. Warner, 2003: "Analysis and Measures of Effectiveness Values for Predictions of the DF-5," in *Final Report on the DISCRETE FURY Test Program*, for DTRA by Applied Research Associates, Inc. (Report # ARA-LR-3.03-001), Shock Physics Division, P.O. Box 5388, Albuquerque, NM 87185, August 2003.

2-12. Warner, S., N. Platt, and J. F. Heagy, 2002: *Explorations to Support the Selection of Trials for HPAC/NARAC Model Comparisons: User-Oriented Measure of Effectiveness Values for the Over-Land Along-Wind Dispersion (OLAD) Field Experiments of September 1997,* IDA Memorandum for DTRA, 24 April 2002.

2-13. a) Platt, N., S. Warner, and J. F. Heagy, 2002: *User-Oriented Measure of Effectiveness Values for Model Predictions of the Transport and Dispersion of Pollutants Inside a Building,* IDA Memorandum for DTRA, 19 June 2002. b) Platt N., S. Warner, and J. F. Heagy, 2002: Application of two-dimensional user-oriented measure of effectiveness to interior building releases. *Sixth Annual George Mason University Transport and Dispersion Modeling Workshop,* Fairfax, VA, July 2002.

2-14. a) Warner, S., J. F. Heagy, N. Platt, D. Larson, G. Sugiyama, J. S. Nasstrom, K. T. Foster, S. Bradley, and G. Bieberbach, 2001: *Evaluation of Transport and Dispersion Models: A Controlled Comparison of Hazard Prediction and Assessment Capability (HPAC) and National Atmospheric Release Advisory Center (NARAC) Predictions,* IDA Paper P-3555. (Available electronically at DTIC STINET ada391555.) b) Warner, S., N. Platt J. F. Heagy, S. Bradley, G. Bieberbach, G. Sugiyama, J. S. Nasstrom, K. T. Foster, and D. Larson, 2001: Model intercomparison with user-oriented measures of effectiveness. *Fifth*

Annual George Mason University Transport and Dispersion Modeling Workshop, Fairfax, VA, July 2001.

2-15. Warner, S., N. Platt, and J. F. Heagy, 2001: *Application of User-Oriented MOE to HPAC Probabilistic Predictions of Prairie Grass Field,* IDA Paper P-3586. (Available electronically at DTIC STINET ada391653.)

2-16. a) Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of Urban 2000 observations and urban HPAC predictions paired in space and time. *Proceedings of the 9th Int'l Conference Harmonisation Within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, 1-4 June 2004, pp 172-176. b) Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of Urban 2000 observations and urban HPAC predictions paired in space and time. *72nd Military Operations Research Society Symposium*, Monterey, California, 22-24 June 2004. c) Warner, S. and N. Platt, 2003: *Analyses in Support of Initial Validation of Urban HPAC: Comparisons to Urban 2000 Observations,* IDA Document D-2870N. (Available from Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.). d) Warner, S. and N. Platt, 2003: Comparisons of Urban 2000 Observations and urban HPAC predictions paired in space and time, part 1: methodology and measures. *Seventh Annual GMU Transport and Dispersion Modeling Workshop*, at George Mason University, 17-19 June 2003. e) Platt, N. and S. Warner, 2003: Comparisons of urban 2000 observations and urban HPAC predictions paired in space and time, part 2: results and conclusions. *Seventh Annual GMU Transport and Dispersion Modeling Workshop*, at George Mason University, 17-19 June 2003. f) Boybeyi, Z., J. C. Chang, S. R. Hanna, P. Franzese, N. Platt, and S. Warner, 2003: *Independent Evaluation of Urban HPAC with the Urban 2000 Field Data.* Prepared for the Defense Threat Reduction Agency.

2-17. Warner, S., N. Platt, J. F. Heagy, J. E. Jordan, and G. Bieberbach, 2005: *Comparisons of transport and dispersion model predictions of the Mock Urban Setting Test (MUST) field experiment.* Institute for Defense Analyses Paper P-4030. (Available from Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, VA 22311-1882.)

2-18. Chang, J. C., P. Franzese, K. Chayantrakom, and S. R. Hanna, 2003: Evaluations of CALPUFF, HPAC, and VLSTRACK with two mesoscale field datasets. *J. Appl. Meteor.,* **42**, 453-466.

2-19. Hanna, S. R., J. C. Chang, R. Britter, and M. Neophytou, 2003b: Overview of model evaluation history and procedures in the atmospheric air quality area. *QNET-CFD Network Newsletter,* **2**, 1-4.

2-20. Boybeyi, Z., N. N. Ahmad, D. P. Bacon, T. J. Dunn, M. S. Hall, P. C. S. Lee, R. A. Sarma, and T. R. Wait, 2001: Evaluation of the operational multiscale environment model with grid adaptivity against the European tracer experiment. *J. Appl. Meteor.,* **40**, 1541-1558.

2-21. Eleveld, H., 2001: Application of a methodology to validate atmospheric dispersion models. *Proc. 7th Int'l Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, Belgirate, Italy, JRC-EI, 19-23.

2-22. ASTM, 2000: *Standard guide for statistical evaluation of atmospheric dispersion model performance. American Society for Testing and Materials.* Designation D 6589-00. (Available from ASTM, 100 Barr Harbor Dr., PO Box C700, West Conshohocken, PA, 19428.)

2-23. Mosca, S., G. Graziani, W. Klug, R. Bellasio, and R. Bianconi, 1998: A statistical methodology for the evaluation of long-range dispersion models: an application to the ETEX exercise. *Atmos. Environ.,* **32** (24), 4307-4324.

2-24. Lee, R. F. and J. S. Irwin, 1997: Improving concentration measures used for evaluating air quality models. *J. Appl. Meteor.,* **36**, 1107-1112.

2-25. Hanna, S. R., J. C. Chang, and D. G. Strimaitis, 1993: Hazardous model evaluation with field observations. *Atmos. Environ.,* **27A** (15), 2265-2285.

2-26. Efron, B. and R. J. Tibshirani, 1993: *An Introduction to the bootstrap*. *Monographs on Statistics and Applied Probability,* No. 57, Chapman and Hall.

2-27. Cytel, 1998: *StatXact 4 for Windows, User Manual*, Cytel Software Copyright.

2-28. Sprent, P., 1998: *Data Driven Statistical Methods.* Chapman and Hall.

2-29. Daniel, W. W., 1997: *Applied Nonparametric Statistics.* Wadsworth.

2-30. Sprent, P. and N. C. Smeeton, 2001: *Applied Nonparametric Statistical Methods*, 3d ed. Chapman and Hall.

2-31. Perneger T. V., 1998: What is wrong with Bonferroni adjustments. *British Medical Journal,* **136**, 1236-1238.

CHAPTER 3

COMPARISONS OF URBAN HPAC PREDICTIONS AND JOINT URBAN 2003 OBSERVATIONS

3. COMPARISONS OF HPAC PREDICTIONS AND JOINT URBAN 2003 OBSERVATIONS

This chapter provides comparisons, results, and discussions associated with the Urban HPAC predictions of JU03. The predictions of five Urban HPAC modes – **UC**, **WM**, **DM**, **DW**, and **MS** – are compared using five different MET options – BAS, BRB, PNA, ACA, and PO7 – to the surface sampler observations and to one another. For each comparison, 30-minute average concentration comparisons are examined in the CBD and separately for all of the arc-based samplers (1, 2, and 4 km arcs). The focus of the comparisons in this chapter is on the following metrics: NAD, NMSE, FB, and the MOE – both average concentration- and threshold-based.

A. COMPARSIONS FOR THE BAS MET OPTION

1. Concentration-Based MOE

a. CBD

Figure 3-1 compares concentration-based MOE values for daytime and nighttime JU03 predictions within the CBD using the BAS MET input option and the five Urban HPAC modes that we examined. In general, model performance varied greatly between the day and night releases. For the day releases, four of the Urban HPAC modes led to under-predictions ("above the diagonal"). The exception is the **MS** mode, which led to a slight over-prediction (similar in magnitude to the **DW** under-prediction). In Figure 3-1, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions. We compared these results using the hypothesis test procedure described in the previous chapter and found that no differences (among the 10 possible comparisons) could be discerned (at the 0.0496 significance level as described previously for the 17 independent daytime releases).

At night, all five modes led to over-predictions with the **MS** mode resulting in the least over-prediction. In terms of having an MOE value closest to the perfect value of (1,1), the nighttime results show that **MS**, **DW**, and **DM** are closer than **UC** and **WM**.

Hypothesis testing revealed significant differences (at the 0.0112 significance level), with the first mode listed being the one with values closer to (1,1) and having less false negative *and* less false positive, for the following six comparisons: **DM**-**UC**, **DW**-**UC**, **MS**-**UC**, **DM**-**WM**, **DW**-**WM**, and **MS**-**WM**. The other four comparisons did not lead to statistically significant improved performance of one mode relative to the other.

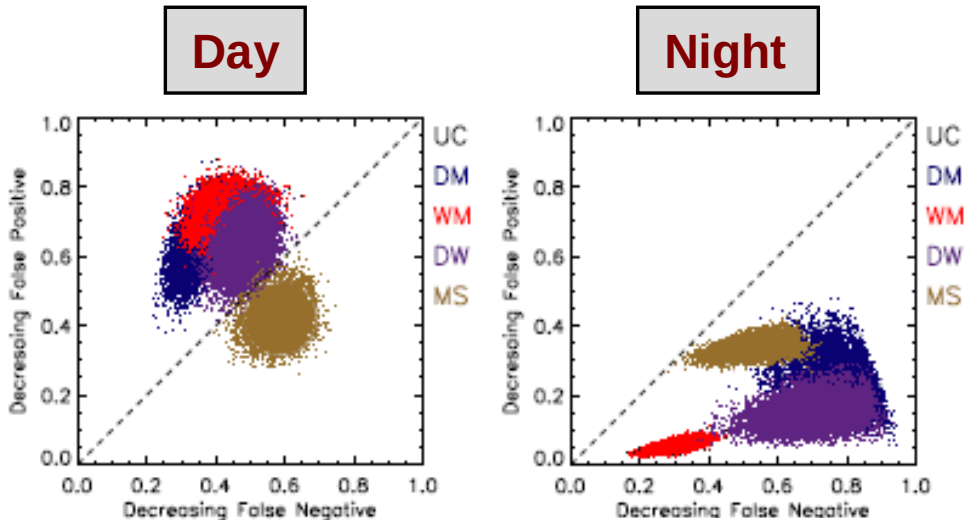


Figure 3-1. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the BAS MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

b. Arcs

Figure 3-2 compares concentration-based MOE values for daytime and nighttime JU03 predictions along the arcs using the BAS MET input option and the five Urban HPAC modes that we examined. For the day releases, all five Urban HPAC modes led to under-predictions. For the night releases, all five Urban HPAC modes led to over-predictions.[1] Although the **MS** predictions result in the cluster that is closest to (1,1) for the day releases – just slightly closer than the **DW** cluster, formal hypothesis testing

[1] In Figure 3-2, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions. In general, as will be seen throughout this chapter, the **UC** and **WM** results were very similar. In fact, the performance of these two modes was so similar that we worried that a software error of some sort may be influencing this result (e.g., perhaps the **WM** mode is not being properly applied within HPAC when run in the manner used here). We continue to work with the model developer to investigate this potential issue.

revealed no significant differences (at least with respect to one set of MOE values having less false positive *and* less false negative as is tested by our procedure) among Urban HPAC modes for daytime JU03 predictions on the arcs (at 1, 2, and 4 km). For the night predictions, hypothesis testing suggested a significant improvement for the **DM** mode relative to the **UC**, **WM**, and **DW** modes. Note that the **MS** mode confidence region is located at similar y-values (similar false positive values) to the **DM** mode, and therefore, our chosen hypothesis test, which demands less false positive *and* less false negative, does not indicate a significant difference for the **MS**-**DM** comparison.

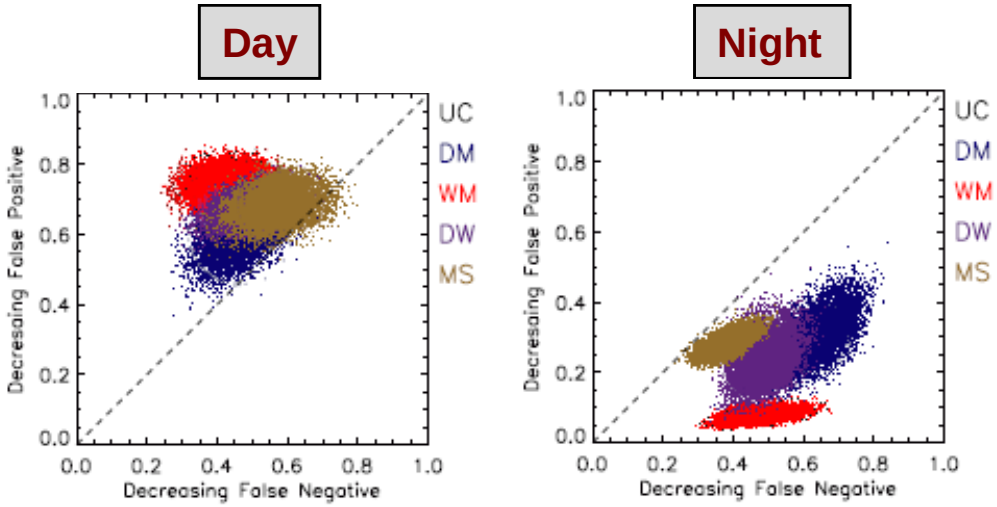


Figure 3-2. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the BAS MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

2. Fractional Bias (FB)

a. CBD

Figure 3-3 compares FB values for daytime and nighttime JU03 predictions within the CBD using the BAS MET input option and the five Urban HPAC modes that we examined. As seen previously with the concentration-based MOE, for the day releases, four of the Urban HPAC modes led to under-predictions (FB < 0). The exception is the **MS** mode, which led to a slight over-prediction (similar in magnitude to the **DW** under-prediction). At night, all five modes led to over-predictions with the **MS** mode resulting in the least over-prediction. Hypothesis test results for FB comparisons

(in the CBD) are shown in Table 3-1.[2] For the day releases, the **MS** FB values are found to be significantly different (p-value < 0.05) from those associated with the **UC**, **DM**, **WM**, and **DW** predictions, and the **DW**-**DM** comparison leads to a significant difference, with **DW** predictions resulting in less under-prediction than **DM**. For the night releases, the **MS** FB values again differ from all other modes – less over-prediction at night – and the **DM**-based predictions resulted in less over-prediction relative to **UC**, **WM**, and **DW**.

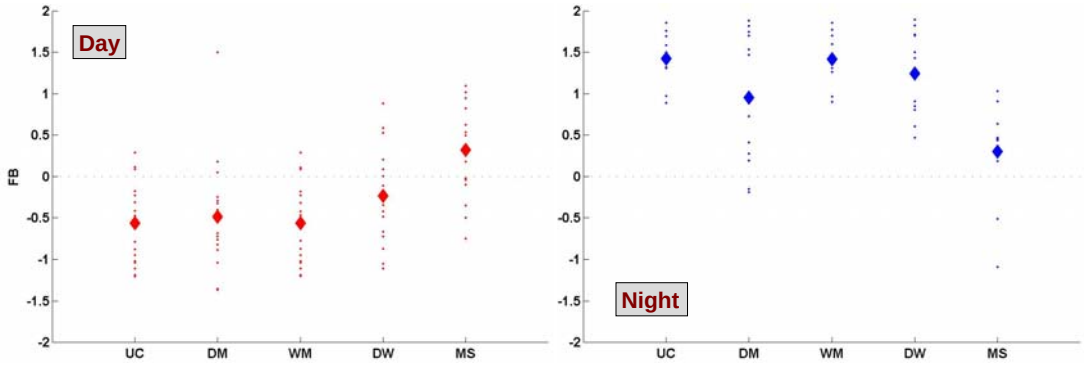


Figure 3-3. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations within the CBD using the BAS MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

a. Arcs

Figure 3-4 compares FB values for daytime and nighttime JU03 predictions on the arcs using the BAS MET input option. The results are reasonably similar to those described for the CBD. Hypothesis test results for FB comparisons (on the arcs) are shown in Table 3-1. For FB comparisons, the absolute value closer to 0.0 corresponds to less under- or over-prediction. The test results shown in Table 3-2 suggest improved FB performance for **MS**, **DM**, and **DW** relative to **UC** and **WM** for both the day and night. In addition, during the day, **MS** predictions led to less under-prediction than **DW** and, at night, **MS** resulted in less over-prediction than **DM** and **DW**.

---

[2] P-values below 0.05 are considered significant for this study (e.g., for testing for differences between one-dimensional metrics) and are highlighted with boldface text in Table 3-1 (and the tables that follow). Text is colored (e.g., purple for **DW**) to indicate to the mode with the better value (e.g., lower NAD value or lower *absolute* FB value). It should be noted that the paired, general permutation procedure used for the hypothesis testing of the one-dimensional metrics (e.g., FB) is quite sensitive to small differences (i.e., able to detect), assuming they are consistent across many releases.

Table 3-1. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the BAS MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.7285	**0.0341**
WM-UC	0.9920	0.4717
DW-UC	0.0839	0.0764
MS-UC	**< 0.0001**	**0.0010**
DM-WM	0.7258	**0.0318**
DM-DW	**0.0016**	**0.0261**
DM-MS	**0.0010**	**0.0073**
DW-WM	0.0842	0.0767
MS-WM	**< 0.0001**	**0.0002**
DW-MS	**0.0072**	**0.0007**

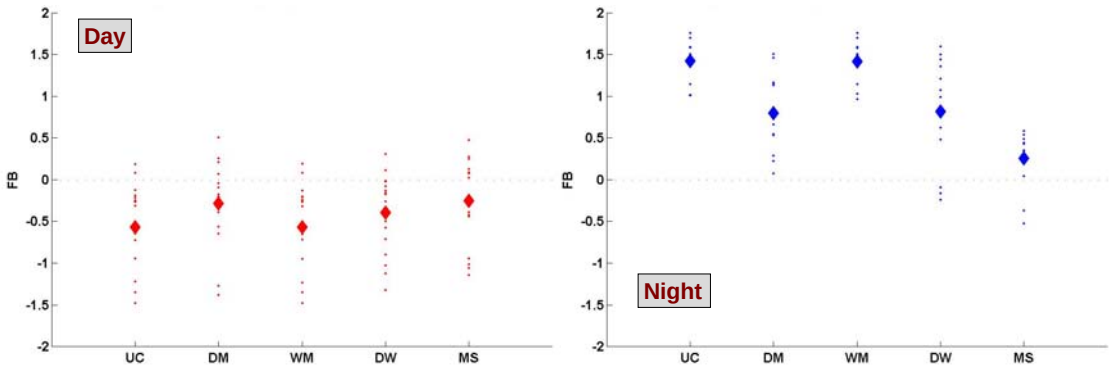


Figure 3-4. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations on the arcs using the BAS MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

Table 3-2. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the BAS MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	**0.0004**	**0.0006**
WM-UC	0.9209	0.3209
DW-UC	**< 0.0001**	**0.0004**
MS-UC	**< 0.0001**	**0.0004**
DM-WM	**0.0002**	**0.0001**
DM-DW	0.1632	0.8523
DM-MS	0.7643	**0.0028**
DW-WM	**< 0.0001**	**0.0003**
MS-WM	**0.0001**	**0.0005**
DW-MS	**0.0009**	**0.0316**

3. Normalized Absolute Difference (NAD)

a. CBD

Figure 3-5 compares NAD values for daytime and nighttime JU03 predictions within the CBD using the BAS MET input option and the five Urban HPAC modes that we examined. Table 3-3 presents the corresponding hypothesis test results. For the day predictions, **UC**, **WM**, and **DW** all result in lower scatter (a smaller NAD value and a better fit with the observations) than **DM**. At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **MS** and **DM** result in improvement over **DW**.

b. Arcs

Figure 3-6 compares NAD values for daytime and nighttime JU03 predictions on the arcs using the BAS MET input option. Hypothesis test results for NAD comparisons (on the arcs) are shown in Table 3-4. These test results suggest:

- For the day predictions, **MS** and **DW** result in lower scatter than **UC**, **WM**, and **DM**. **MS**-based predictions led to lower scatter than **DW**-based predictions.

- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **DM** results in improvement over **DW**. These results are quite similar to those found for NAD in the CBD.

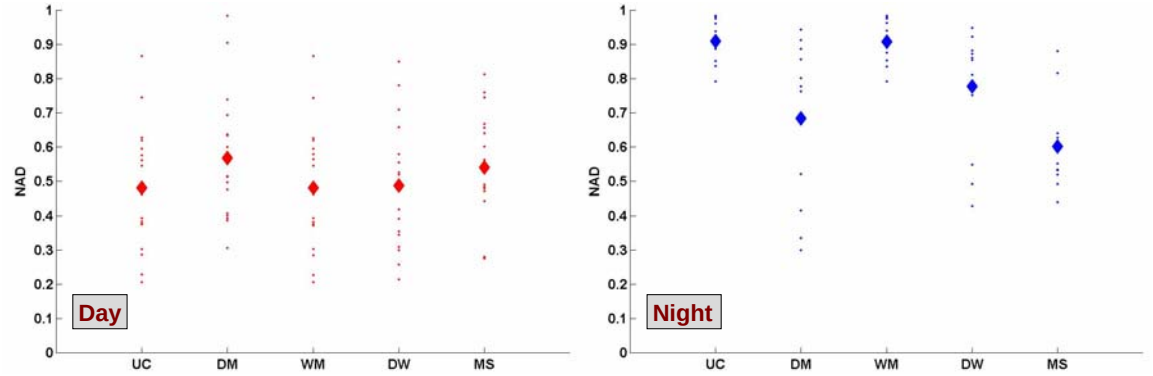


Figure 3-5. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations within the CBD using the BAS MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-3. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the BAS MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	**0.0038**	**0.0005**
WM-UC	0.2910	0.2088
DW-UC	0.7424	**0.0007**
MS-UC	0.2816	**0.0006**
DM-WM	**0.0035**	**0.0006**
DM-DW	**< 0.0001**	**0.0127**
DM-MS	0.5790	0.2569
DW-WM	0.7217	**0.0007**
MS-WM	0.2794	**0.0006**
DW-MS	0.3344	**0.0076**

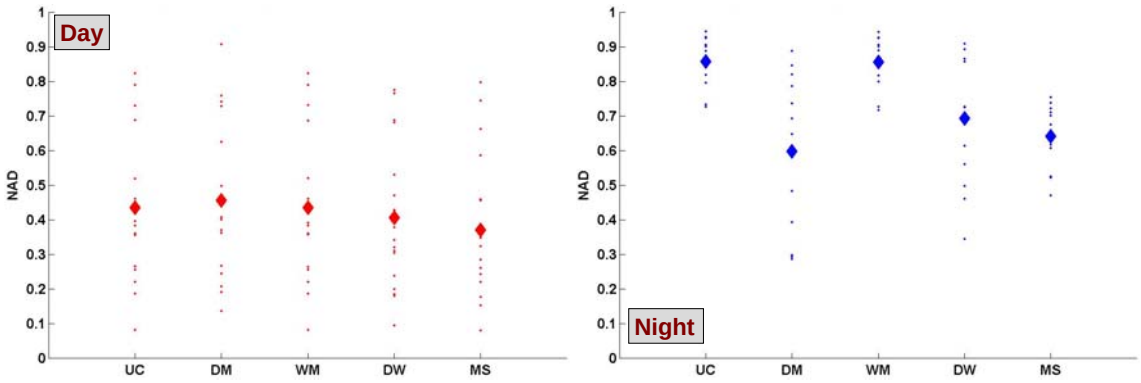


Figure 3-6. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations on the arcs using the BAS MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-4. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the BAS MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.1207	**0.0006**
WM-UC	0.4588	0.5309
DW-UC	**0.0011**	**0.0004**
MS-UC	**0.0019**	**0.0001**
DM-WM	0.1238	**0.0007**
DM-DW	**0.0004**	**0.0013**
DM-MS	**0.0014**	0.4058
DW-WM	**0.0010**	**0.0007**
MS-WM	**0.0022**	**0.0007**
DW-MS	**0.0292**	0.2063

4. Normalized Mean Square Error (NMSE)

a. CBD

Figure 3-7 compares NMSE values for daytime and nighttime JU03 predictions within the CBD using the BAS MET input option. Table 3-5 presents the hypothesis test results that compare NMSE values for daytime and nighttime JU03 predictions within the CBD using the BAS MET input option. The significant results are summarized below.

- For the day predictions, **DW** results in lower scatter (smaller value of NMSE) than **DM**, and **UC** results in lower scatter than **WM**. While the **UC-WM** difference is detectable with our paired hypothesis testing procedures, it is not considered particularly important since the magnitude of the differences is so small.
- For the night predictions, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **MS** results in improvement over **DM**. These results are similar to those found for NAD comparisons at night in the CBD.

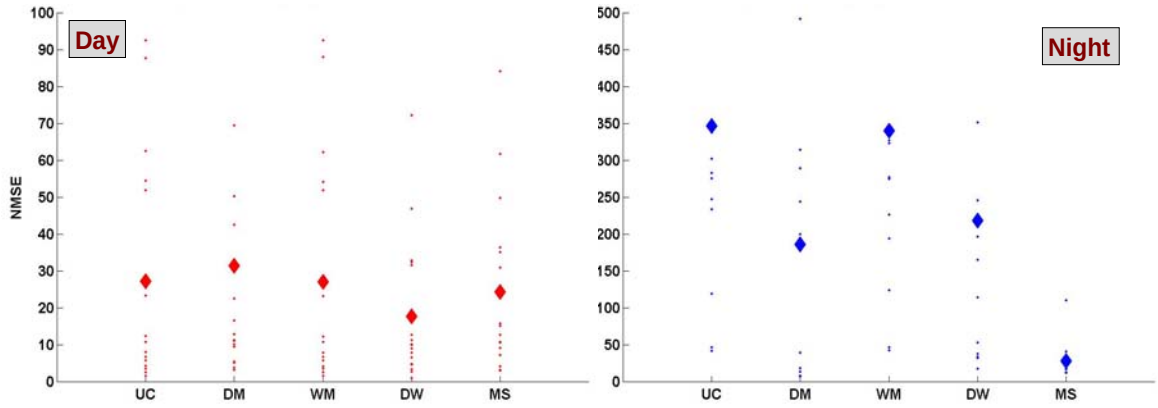


Figure 3-7. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations within the CBD using the BAS MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. Some of the larger values are off the depicted scale. The larger colored diamonds correspond to the average NMSE value.

Table 3-5. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the BAS MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.6989	**0.0019**
WM-UC	**0.0404**	0.3738
DW-UC	0.0980	**0.0012**
MS-UC	0.7830	**0.0017**
DM-WM	0.6948	**0.0026**
DM-DW	**0.0002**	0.1473
DM-MS	0.4962	**0.0331**
DW-WM	0.0975	**0.0021**
MS-WM	0.7771	**0.0006**
DW-MS	0.3661	**0.0093**

c. Arcs

Figure 3-8 compares NMSE values for daytime and nighttime JU03 predictions on the arcs using the BAS MET input option. Hypothesis test results for NMSE comparisons (on the arcs) are shown in Table 3-6. These test results suggest:

- For the day predictions, **MS** and **DW** result in lower scatter than **UC** and **WM**, and **MS**-based predictions led to lower scatter than **DW**-based predictions. This result is very similar to that found for NAD during the day on the arcs.
- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **MS**- and **DM**-based predictions resulted in improvement over **DW**. These results are quite similar to those found for NAD on the arcs at night.

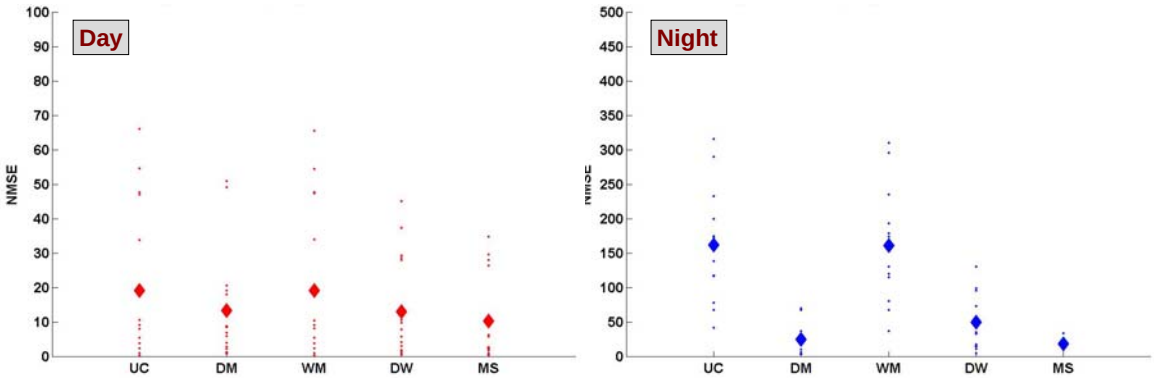


Figure 3-8. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations on the arcs using the BAS MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. The larger colored diamonds correspond to the average NMSE value.

Table 3-6. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the BAS MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.0954	**0.0004**
WM-UC	0.8176	0.5504
DW-UC	**0.0005**	**0.0010**
MS-UC	**0.0006**	**0.0008**
DM-WM	0.0904	**0.0006**
DM-DW	0.8205	**0.0014**
DM-MS	0.1038	0.4040
DW-WM	**0.0004**	**0.0005**
MS-WM	**0.0003**	**0.0006**
DW-MS	**0.0097**	**0.0132**

5. Threshold-Based MOE

a. CBD

Figure 3-9 compares threshold-based MOE values for daytime and nighttime JU03 predictions within the CBD using the BAS MET input option and the five Urban HPAC modes that we examined. Two thresholds are considered, 25 and 250 ppt, which correspond to 5 and 50 × the estimated SF_6 background. For the day releases, all five Urban HPAC modes resulted in over-predictions at the lower threshold – that is, more samplers were predicted to exceed 25 ppt than were observed to exceed 25 ppt. For the 250 ppt threshold, the daytime predictions vary from slight under- to slight over-predictions (of the number of samplers exceeding 250 ppt).[3] For the nighttime predictions, the **MS**, **DM**, and **DW** modes led to over-predictions, albeit very slight for **MS** at 250 ppt, and the **UC** and **WM** modes resulted in under-predictions for both thresholds examined in Figure 3-9. The overall improvement in MOE value for the threshold-based computations relative to the concentration-based values (Figure 3-1) is notable (i.e., the threshold-based MOE values are closer to the perfect value of (1,1) than the concentration-based MOE values, particularly at night). As previous studies (of nighttime releases) have found [Refs. 3-1 and 3-2], predictions of the locations (and times) of exceeding low concentration thresholds are greatly improved relative to predictions of actual concentration values in time and space.

We compared these results using the hypothesis test procedure for the MOE. No statistically significant "improvements" (less false positive *and* less false negative, which is what we test for) were found. Clearly there are differences (as described above) between relative under- and over-prediction, but these are not necessarily defined as improvements on their own, e.g., some modes result in predictions that imply less false negative but similar false positive, and our hypothesis test procedure for the MOE requires that both false positive and false negative be improved significantly.

---

[3] In Figure 3-9, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions.

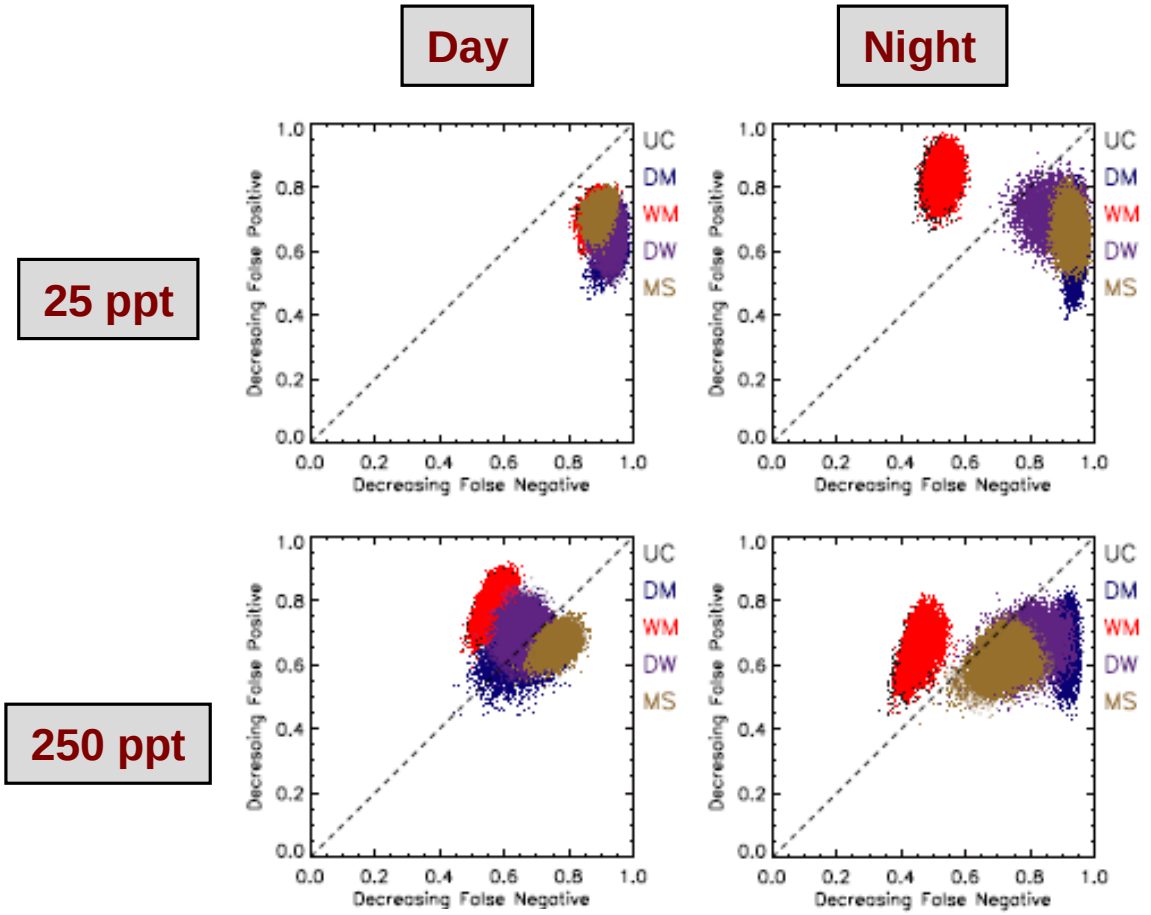


Figure 3-9. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the BAS MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

b. Arcs

Figure 3-10 compares threshold-based MOE values for daytime and nighttime JU03 predictions along the arcs using the BAS MET input option. For the day releases, and the 25 ppt threshold, all five Urban HPAC modes led to MOE values that cluster near the diagonal. At 250 ppt, all five modes resulted in slight under-predictions. At night, all five Urban HPAC modes resulted in over-predictions for both thresholds. The long elliptical nature of the confidence region clusters indicates that, for some of the releases, the wind direction and/or wind speed was not well matched to what was experienced by the plume, at least by the time the plume arrived at the 1, 2, and 4 km arcs. Such a miss of the direction/speed leads to the model predicting material at locations where little is

observed – false positive – and predicting little material at locations where larger concentrations are observed – false negative. Therefore, both the false positive *and* the false negative increase and the MOE value moves down the diagonal. Where this phenomena occurs for several of the releases being considered (at least at the arcs in this case), the confidence region cluster takes on the shape shown in Figure 3-10. This phenomenon is particularly prevalent on the arcs at the higher threshold because less material reaches these samplers, and hence, fewer samplers actually exceed the threshold. Then, relatively small wind direction errors result in the prediction missing the few samplers that actually did exceed the threshold, at least on some of the releases.

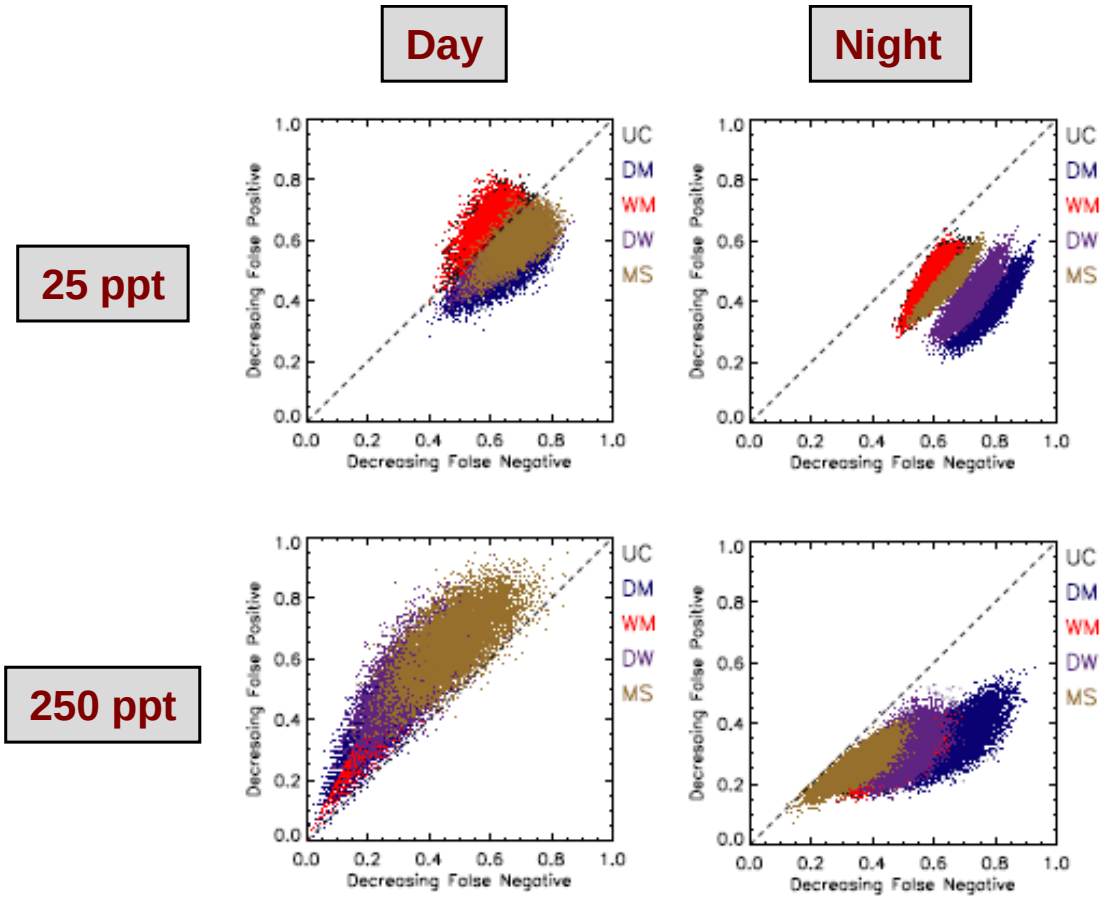


Figure 3-10. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the BAS MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

Hypothesis test results for threshold-based MOE comparisons (on the arcs) resulted in the following conclusions:

- For the day predictions, **MS**-based predictions led to improved MOE values relative to **UC** and **WM** for the 250 ppt threshold.
- At night, the **DM** mode resulted in improved MOE values for the 250 ppt threshold when compared to **UC**, **WM**, **DW**, and **MS**.

6. Summary for BAS MET Input Option Results

Based on the above discussions and consideration of the comparisons of the five metrics (especially the scatter metrics – NAD, NMSE, and concentration-based MOE), the qualitative results shown in Table 3-7 appear reasonable. For the daytime releases and predictions within the CBD, the hypothesis test results led to mixed conclusions with no single Urban HPAC mode consistently resulting in improved performance. The "**DW**/**DM**" for the Day, CBD in Table 3-7, implies that for two of the three scatter-related metrics, **DW** outperformed (i.e., "less scatter" for two of the three scatter-related metrics) **DM**. For the predictions on the arcs during the day **MS** and **DW** outperformed **UC** and **WM**. For all night predictions (CBD and arcs), the **MS**, **DM**, and **DW** modes resulted in improved performance relative to the **UC** and **WM** modes and additionally on the arcs, **DM** outperformed **DW**.

These results are especially important because the BAS MET input option, which served as a surrogate for use of the DTRA MET server, is representative of an operational capability with respect to the availability of meteorological information. That is, using the DTRA MET server to obtain nearby wind observations – perhaps just a few hours after the release – to create plume predictions, as was the case for the BAS MET input option, corresponds to a reasonably realistic operational usage of Urban HPAC.

Table 3-7. Urban HPAC Modes (BAS MET) That Consistently, Across at Least Two of the Three Scatter-Related Metrics, Led to Improved Predictive Performance of JU03 Releases

Surface Sampler Region	Day	Night
CBD	**DW**/**DM**	(**MS**,**DM**,**DW**)/(**UC**,**WM**)
Arcs	(**MS**,**DW**)/(**UC**,**WM**)	(**MS**,**DM**,**DW**)/(**UC**,**WM**) and **DM**/**DW**

B. COMPARSIONS FOR THE BRB MET OPTION

1. Concentration-Based MOE

a. CBD

Figure 3-11 compares concentration-based MOE values for daytime and nighttime JU03 predictions within the CBD using the BRB MET input option and the five Urban HPAC modes that we examined. For the day releases, four of the Urban HPAC modes led to under-predictions ("above the diagonal"). The exception is the **MS** mode, which predicted about the right amount of material at the surface samplers.[4] Using the hypothesis test procedure described in the previous chapter, we found that no differences (among the 10 possible comparisons) could be discerned (at the 0.0496 significance level as described previously for the 17 independent daytime releases).

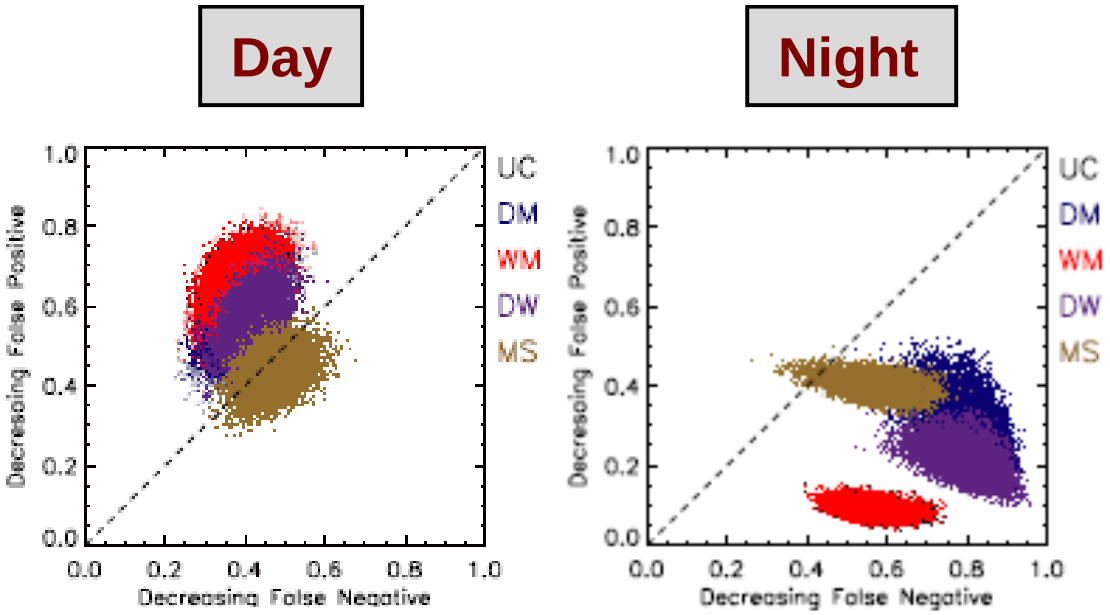


Figure 3-11. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the BRB MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

---

[4] In Figure 3-11, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions.

At night, all five modes led to over-predictions with the **MS** mode resulting in the least over-prediction. In terms of having an MOE value closest to the perfect value of (1,1), the nighttime results show that **MS**, **DW**, and **DM** are closer than **UC** and **WM**. Hypothesis testing revealed significant differences (at the 0.0112 significance level), with the first mode listed being the one with values closer to (1,1) and having less false negative *and* less false positive, for the following four comparisons: **DM**-**UC**, **DW**-**UC**, **DM**-**WM**, and **DW**-**WM**.

b. Arcs

Figure 3-12 compares concentration-based MOE values for daytime and nighttime JU03 predictions along the arcs using the BRB MET input option. For the day releases, all five Urban HPAC modes led to under-predictions, albeit slight in some cases. For the night releases, four of the five Urban HPAC modes led to over-predictions with the **MS** predictions being the exception (as the MOE confidence region straddles the diagonal). Hypothesis testing revealed improvements in the MOE values associated with daytime **DW**, **UC**, and **WM** predictions on the arcs relative to the **DM** predictions. For the night predictions, no significant improvements ("less false negative and less false positive") were suggested by the hypothesis testing using the concentration-based MOE and considering the samplers on the arcs.

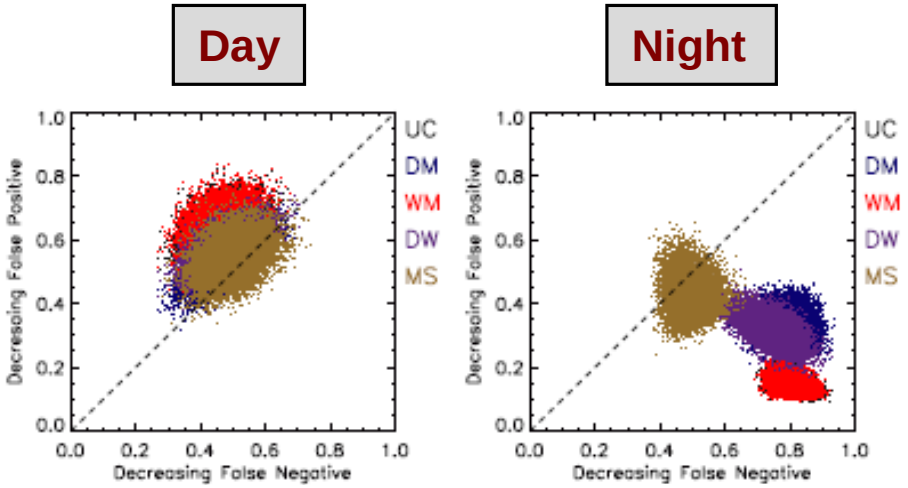


Figure 3-12. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the BRB MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

2. Fractional Bias (FB)

a. CBD

Figure 3-13 compares FB values for daytime and nighttime JU03 predictions within the CBD using the BRB MET input option. As seen previously with the concentration-based MOE, for the day releases, four of the Urban HPAC modes led to under-predictions (FB < 0). The exception is the **MS** mode, which led to a slight over-prediction. At night, all five modes led to over-predictions with the **MS** mode resulting in only a minor over-prediction. Hypothesis test results for FB comparisons (in the CBD) are shown in Table 3-8. For the day releases, the **MS** FB values are found to be significantly different (p-value < 0.05) from those associated with the **UC**, **DM**, **WM**, and **DW** predictions, and the **DW**-**WM** and **UC**-**WM** comparisons resulted in significant differences,[5] with the **DW** and **UC** predictions resulting in less under-prediction than **WM**. For the night releases, the **MS** FB values again differ from all other modes – less over-prediction at night – and the **DM**-based predictions resulted in less over-prediction relative to **UC**, **WM**, and **DW**. In addition, at night and for the CBD, the **DW**-based FB values suggested less over-prediction relative to **UC** and **WM** based predictions.

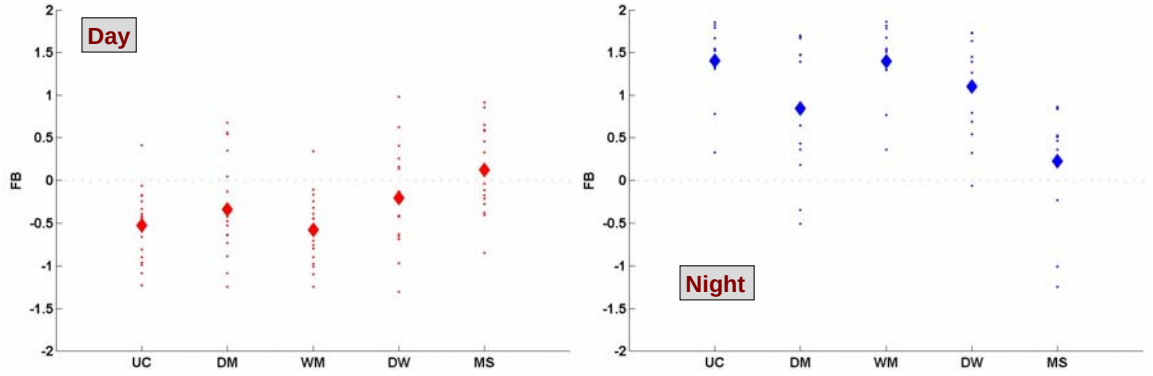


Figure 3-13. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations within the CBD using the BRB MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

---

[5] While this **UC**-**WM** difference is detectable with our paired hypothesis testing procedures, it is not considered particularly important since the magnitude of the difference is so small.

Table 3-8. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the BRB MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.2499	**0.0003**
WM-UC	**0.0051**	0.2960
DW-UC	0.0681	**0.0026**
MS-UC	**< 0.0001**	**0.0004**
DM-WM	0.1638	**0.0021**
DM-DW	0.1582	**0.0039**
DM-MS	**0.0149**	**0.0164**
DW-WM	**0.0474**	**0.0044**
MS-WM	**< 0.0001**	**0.0007**
DW-MS	**0.0294**	**0.0027**

b. Arcs

Figure 3-14 compares FB values for daytime and nighttime JU03 predictions on the arcs using the BRB MET input option. Hypothesis test results for FB comparisons on the arcs are shown in Table 3-9. For FB comparisons, the absolute value closer to 0.0 corresponds to less under- or over-prediction. The test results shown in Table 3-9 suggest improved FB performance for **MS**, **DM**, and **DW** relative to **UC** and **WM** for both the day and night. In addition, during the day, **MS** predictions led to less under-prediction than **DW,** and **UC** resulted in slightly less under-prediction than **WM** (as judged by our paired hypothesis testing). At night, **MS** resulted in less over-prediction than **DM** and **DW**. These results are very similar to those described for the BAS MET input option.

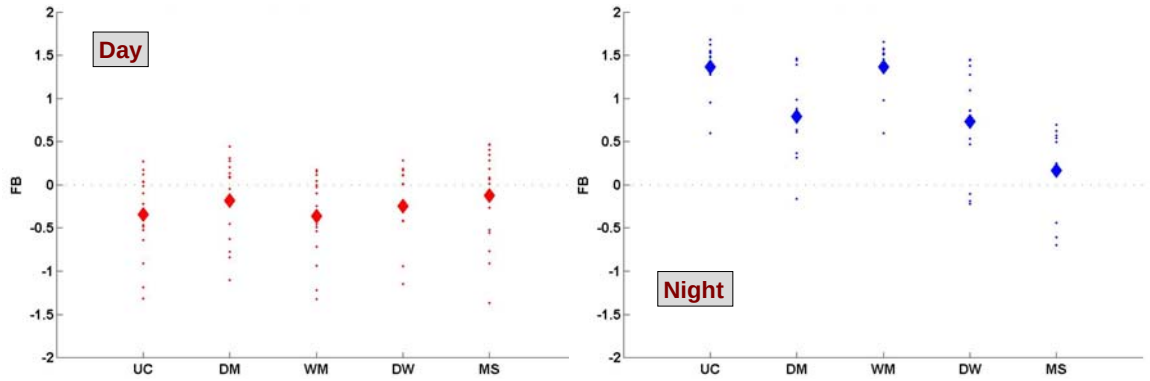


Figure 3-14. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations on the arcs using the BRB MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

Table 3-9. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the BRB MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	< 0.0001	0.0019
WM-UC	0.0017	0.7997
DW-UC	0.0015	0.0023
MS-UC	0.0001	0.0002
DM-WM	< 0.0001	0.0010
DM-DW	0.1186	0.7112
DM-MS	0.2673	0.0010
DW-WM	0.0003	0.0011
MS-WM	< 0.0001	0.0004
DW-MS	0.0167	0.0336

3. Normalized Absolute Difference (NAD)

a. CBD

Figure 3-15 compares NAD values for daytime and nighttime JU03 predictions within the CBD using the BRB MET input option and the five Urban HPAC modes that we examined. Table 3-10 presents the corresponding hypothesis test results. For the day predictions, **UC**, **WM**, and **DW** all result in lower scatter than **DM**, and **WM** led to less scatter than **UC**. At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **DM** results in improvement over **DW**. These results are very similar to those found for the BAS MET input option.

b. Arcs

Figure 3-16 compares NAD values for daytime and nighttime JU03 predictions on the arcs using the BRB MET input option. Hypothesis test results for NAD comparisons (on the arcs) are shown in Table 3-11. These test results suggest:

- For the day predictions, **MS** results in lower scatter than **UC** and **WM**.
- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**. These results are quite similar to those found for NAD in the CBD (and for the arcs and the BAS MET input option).

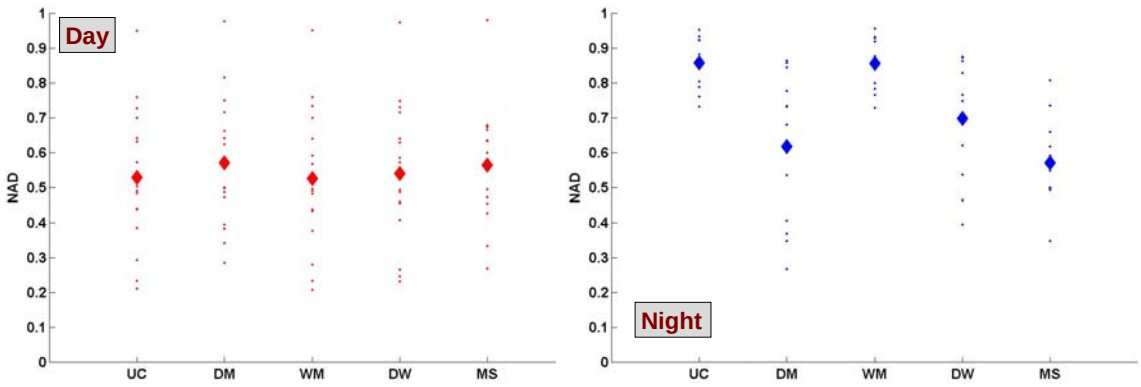


Figure 3-15. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations within the CBD using the BRB MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-10. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the BRB MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	**0.0098**	**0.0006**
WM-UC	**0.0326**	0.1647
DW-UC	0.4028	**0.0004**
MS-UC	0.4087	**0.0018**
DM-WM	**0.0046**	**0.0002**
DM-DW	**0.0321**	**0.0047**
DM-MS	0.8954	0.5503
DW-WM	0.1939	**0.0007**
MS-WM	0.3596	**0.0014**
DW-MS	0.5787	0.0865

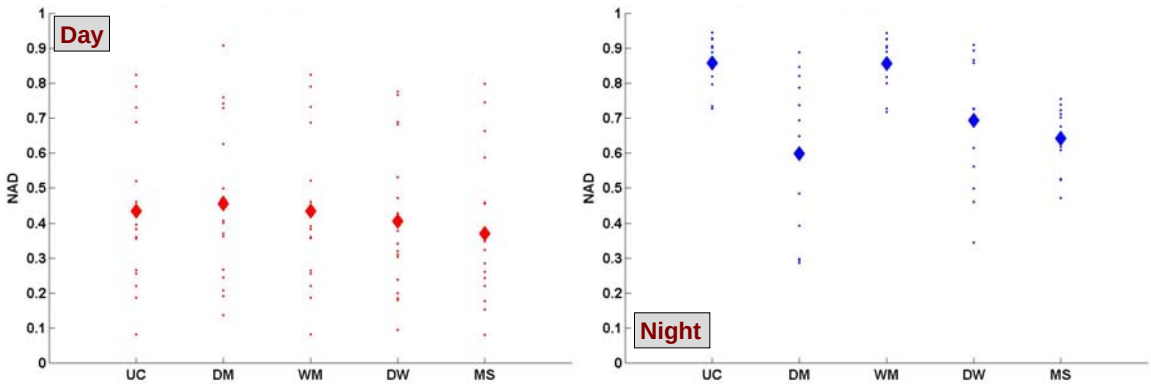


Figure 3-16. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations on the arcs using the BRB MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-11. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the BRB MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.1365	**0.0047**
WM-UC	0.8186	0.9853
DW-UC	0.4661	**0.0010**
MS-UC	**0.0312**	**0.0004**
DM-WM	0.1042	**0.0038**
DM-DW	0.1378	0.1172
DM-MS	0.9360	0.9697
DW-WM	0.1155	**0.0010**
MS-WM	**0.0243**	**0.0005**
DW-MS	0.2605	0.3652

4. Normalized Mean Square Error (NMSE)

a. CBD

Figure 3-17 compares NMSE values for daytime and nighttime JU03 predictions within the CBD using the BRB MET input option. Table 3-12 presents the hypothesis test results that compare NMSE values for daytime and nighttime JU03 predictions within the CBD using the BRB MET input option. We found that no differences (among the 10 possible comparisons) could be discerned (at the 0.05 significance level). For the night predictions, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **MS** results in improvement over **DW**. These results are similar to those found for NAD comparisons at night in the CBD (and to the corresponding BAS MET input option night results).

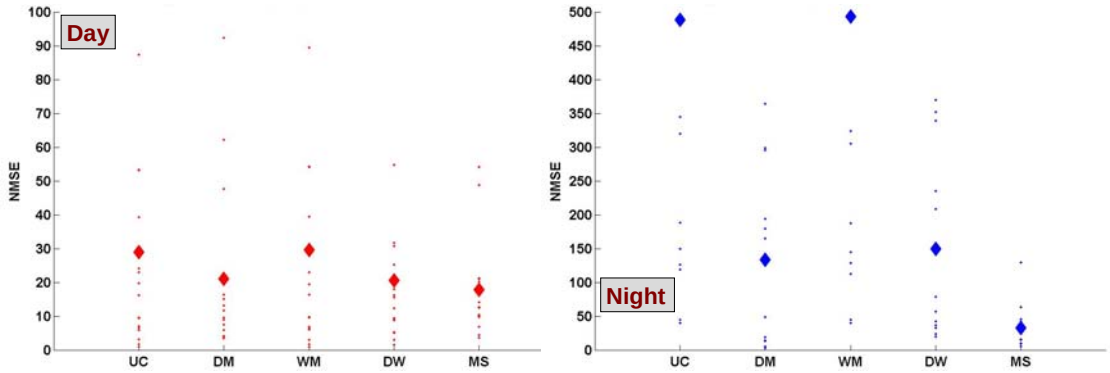


Figure 3-17. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations within the CBD using the BRB MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. Some of the larger values are off the depicted scale. The larger colored diamonds correspond to the average NMSE value.

Table 3-12. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the BRB MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.1983	**0.0009**
WM-UC	0.0518	0.9857
DW-UC	0.2308	**0.0017**
MS-UC	0.2224	**0.0020**
DM-WM	0.1537	**0.0007**
DM-DW	0.7938	0.1590
DM-MS	0.5528	0.0504
DW-WM	0.1735	**0.0020**
MS-WM	0.1975	**0.0013**
DW-MS	0.5950	**0.0419**

b. Arcs

Figure 3-18 compares NMSE values for daytime and nighttime JU03 predictions on the arcs using the BRB MET input option. Hypothesis test results for NMSE comparisons (on the arcs) are shown in Table 3-13. These test results suggest:

- No significant differences for daytime predictions.
- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**, and **MS**-based predictions resulted in improvement over **DW**. These results are quite similar to those found previously for the CBD and on the arcs for NAD.

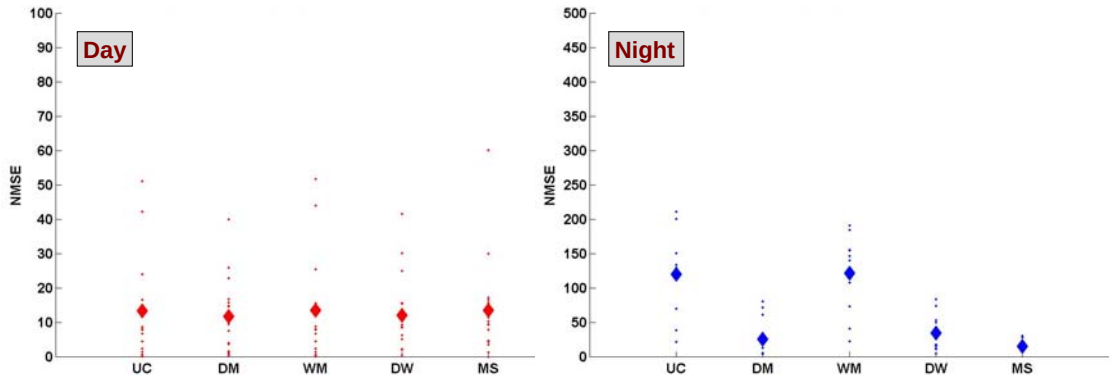


Figure 3-18. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations on the arcs using the BRB MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. The larger colored diamonds correspond to the average NMSE value.

5. Threshold-Based MOE

a. CBD

Figure 3-19 compares threshold-based MOE values for daytime and nighttime JU03 predictions within the CBD using the BRB MET input option and the five Urban HPAC modes that we examined. Two thresholds are considered as previously discussed. For the day releases, all five Urban HPAC modes resulted in over-predictions at the lower threshold – that is, more samplers were predicted to exceed 25 ppt than were observed to exceed 25 ppt. For the 250 ppt threshold, the daytime predictions vary from

slight under- to slight over-predictions (of the number of samplers exceeding 250 ppt).[6] For the nighttime predictions, the **MS**, **DM**, and **DW** modes led to over-predictions, albeit very slight for **MS** at 250 ppt, and the **UC** and **WM** modes resulted in under-predictions for both thresholds examined in Figure 3-19. The overall improvement in MOE value for the threshold-based computations relative to the concentration-based values (Figure 3-11) is notable (i.e., the threshold-based MOE values are closer to the perfect value of (1,1) than the concentration-based MOE values).

Table 3-13. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the BRB MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.2560	**0.0004**
WM-UC	0.2949	0.7585
DW-UC	0.2213	**0.0004**
MS-UC	0.9489	**0.0005**
DM-WM	0.2462	**0.0005**
DM-DW	0.6550	0.1812
DM-MS	0.3201	0.1963
DW-WM	0.1941	**0.0005**
MS-WM	0.9377	**0.0000**
DW-MS	0.3721	**0.0219**

Using the hypothesis test procedure for the MOE, we found the following significant differences:

- During the day and at the 25 ppt threshold, **UC** and **WM** were nearer the perfect (1,1) value than **MS**. This small, but statistically significant movement can be seen in Figure 3-19.

[6] In Figure 3-19, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions.

- At night, the **DW** predictions led to an improved 250 ppt threshold-MOE value relative to **MS**.

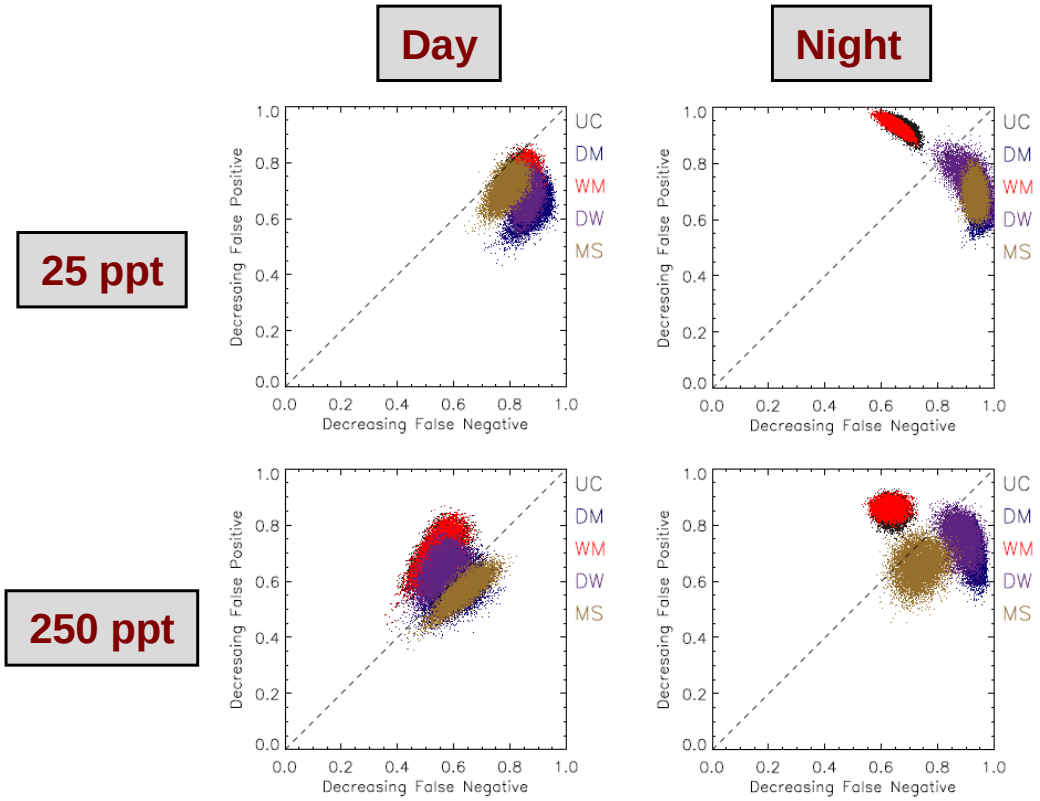


Figure 3-19. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the BRB MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

b. Arcs

Figure 3-20 compares threshold-based MOE values for daytime and nighttime JU03 predictions along the arcs using the BRB MET input option. For the day releases, and the 25 ppt threshold, all five Urban HPAC modes led to MOE values that cluster near the diagonal but correspond to mild over-predictions. At 250 ppt, four of the modes resulted in moderate under-predictions, with **MS** being the exception and showing little bias. At night, all five Urban HPAC modes resulted in over-predictions for both thresholds (albeit very slight for **MS** at the 250 ppt threshold). The long elliptical nature of the confidence region clusters for the daytime data indicates that, for some of the releases, the wind direction and or wind speed, at least by the time the plume arrives at

the 1, 2, and 4 km arcs, was not well matched to what was experienced by the plume (as previously discussed for BAS MET input option). The only significant hypothesis test result for the arc-based comparisons is for the daytime **DW**-**MS** combination; with the **DW**-based 25 ppt threshold-MOE value being somewhat closer to (1,1) than the associated **MS** MOE value. In this case, the actual **MS** MOE value, the average for 17 daytime releases, is closer to (1,1) than the corresponding **DW** value. However, 11 individual-release **DW** values are closer to (1,1) than the corresponding **MS** value, and thus we conclude (by this hypothesis test) that the **DW**-based predictions are "closer" to (1,1). Substantial improvements in performance for **MS** relative to **DW** was associated with just two releases – IOP 2, Releases 2 and 3 – which greatly influenced the above result.

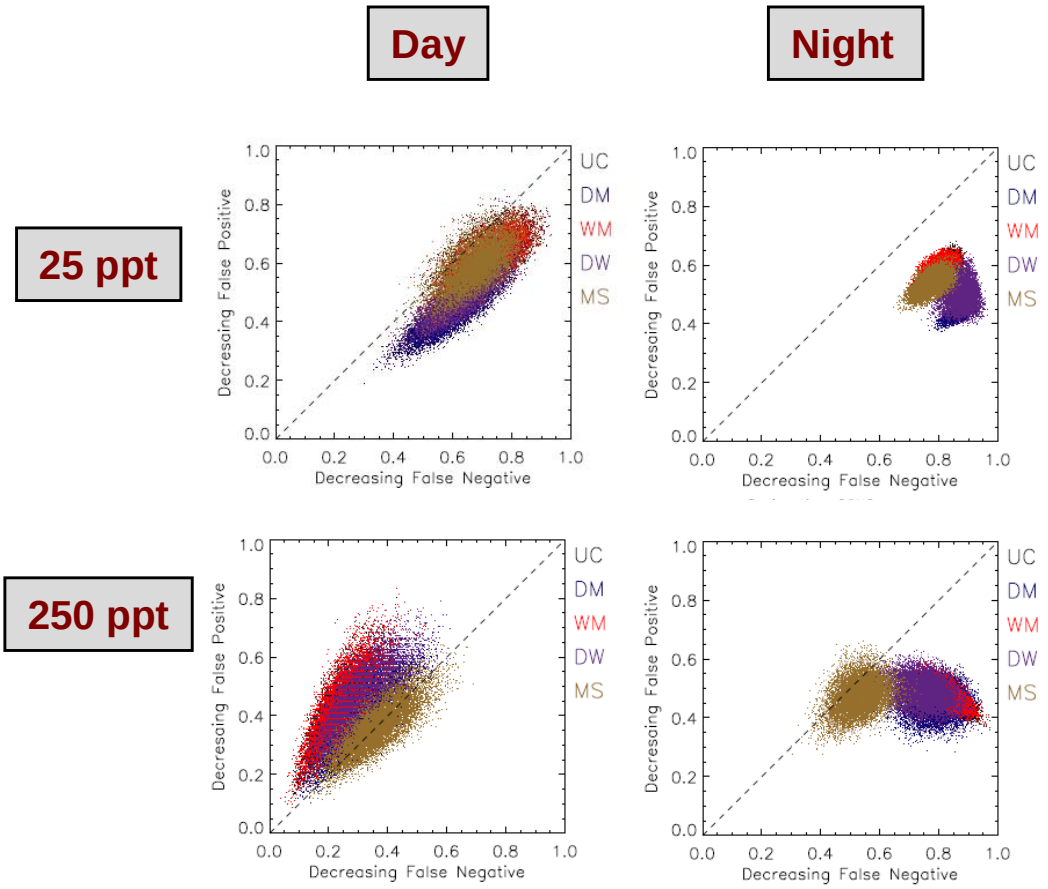


Figure 3-20. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the BRB MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

6. Summary for BRB MET Input Option Results

Based on the above discussions and consideration of the comparisons of the five metrics, particularly the measures of scatter (NAD, NMSE, and concentration-based MOE), the qualitative results shown in Table 3-14 appear reasonable. The improvements in nighttime predictions associated with the inclusion of **MS**, **DM**, and **DW** are similar to what was found for the BAS MET input option. These results can be considered particularly relevant if one considers the inclusion of BRB-like MET input (i.e., numerical weather assimilation) as a reasonably realistic operational usage of Urban HPAC in the aftermath (a few hours) of an event.

Table 3-14. Urban HPAC Modes (BRB MET) That Consistently, Across at Least Two of the Three Scatter-Related Metrics, Led to Improved Predictive Performance of JU03 Releases

Surface Sampler Region	Day	Night
CBD	mixed	(**MS**,**DM**,**DW**)/(**UC**,**WM**)
Arcs	mixed	(**MS**,**DM**,**DW**)/(**UC**,**WM**)

C. COMPARSIONS FOR THE PNA MET OPTION

1. Concentration-Based MOE

a. CBD

Figure 3-21 compares concentration-based MOE values for daytime and nighttime JU03 predictions within the CBD using the PNA MET input option. For the day releases, four of the Urban HPAC modes led to under-predictions ("above the diagonal"). The exception is the **MS** mode, which resulted in an over-prediction.[7] With respect to comparisons, we used hypothesis test procedures to search for sets of predictions that exhibited significantly less false positive *and* significantly less false negative. No significant differences were discerned and for the day or night releases.

b. Arcs

Figure 3-22 compares concentration-based MOE values for daytime and nighttime JU03 predictions along the arcs using the PNA MET input option. For the day

---

7 In Figure 3-21, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions.

releases, all five Urban HPAC modes led to under-predictions, albeit slight in some cases. For the night releases, three of the five Urban HPAC modes led to over-predictions (**MS**, **UC**, and **WM**). Hypothesis testing revealed no significant differences (in terms of false positive *and* false negative) for the daytime release predictions. At night, the **MS** mode led to an improvement relative to the **UC** mode.

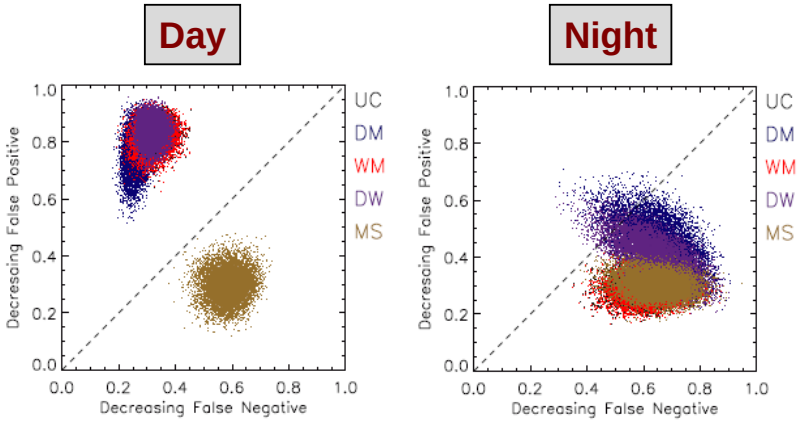


Figure 3-21. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the PNA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

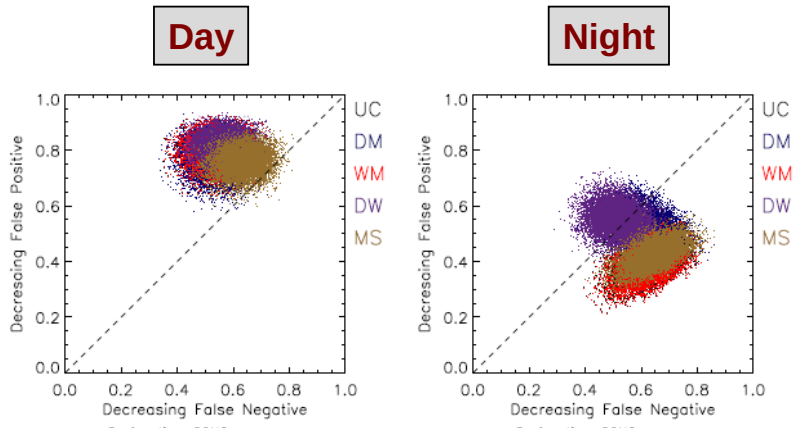


Figure 3-22. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the PNA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

2. Fractional Bias (FB)

a. CBD

Figure 3-23 compares FB values for daytime and nighttime JU03 predictions within the CBD using the PNA MET input option. As seen previously with the concentration-based MOE, for the day releases, four of the Urban HPAC modes led to under-predictions (FB < 0). The exception is the **MS** mode, which led to a significant over-prediction, unlike the BAS and BRB MET cases in which the **MS** mode led to only a slight over-prediction during the day in the CBD. At night, all five modes led to over-predictions. Hypothesis test results for FB comparisons (in the CBD) are shown in Table 3-15. The followings significant findings are obtained:

- For the day releases, the **MS** FB values are found to be significantly different (over-predictions) from those associated with the **UC**, **DM**, **WM**, and **DW** predictions (all under-predictions).
- For the nighttime releases and predictions, the **DM** mode resulted in less over-prediction relative to **UC** and **WM**.

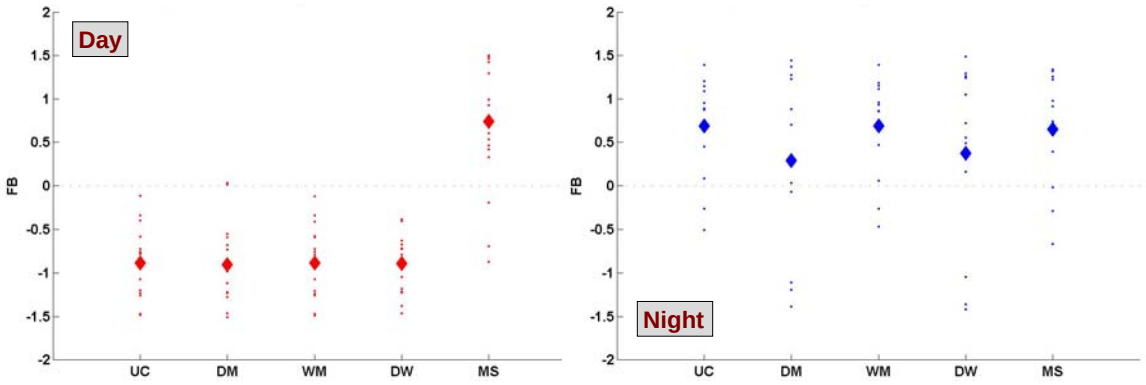


Figure 3-23. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations within the CBD using the PNA MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

c. Arcs

Figure 3-24 compares FB values for daytime and nighttime JU03 predictions on the arcs using the PNA MET input option. Hypothesis test results for FB comparisons on the arcs are shown in Table 3-16 and suggest the following:

Table 3-15. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the PNA MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.9044	**0.0431**
WM-UC	0.6443	0.9099
DW-UC	0.9702	0.0845
MS-UC	**< 0.0001**	0.7867
DM-WM	0.9060	**0.0382**
DM-DW	0.8468	0.3018
DM-MS	**< 0.0001**	0.2878
DW-WM	0.9732	0.0830
MS-WM	**< 0.0001**	0.8000
DW-MS	**0.0001**	0.3809

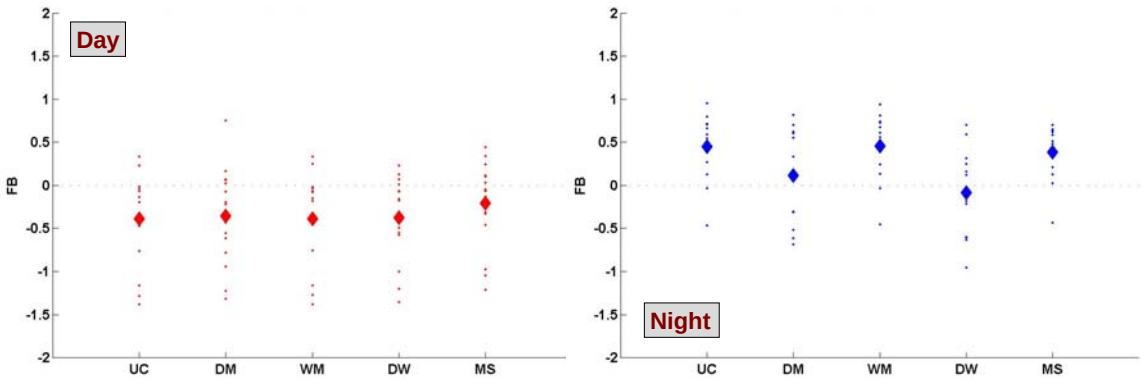


Figure 3-24. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations on the arcs using the PNA MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

- For the day releases, the **MS** FB values are found to be significantly different (and improved, i.e., less under-prediction) from those associated with the **UC**, **WM**, and **DW** predictions.

- At night, the **DM** and **DW** modes resulted in less over-prediction relative to **UC**, **WM**, and **MS**. Also, the FB values associated with the **DW** (FB = -0.086) mode predictions are found to be different from those of **DM** (FB = 0.115), although this is obviously a small (but detectable by our tests) difference.

Table 3-16. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the PNA MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.7845	**0.0091**
WM-UC	0.0610	0.1077
DW-UC	0.6899	**0.0017**
MS-UC	**0.0001**	0.5053
DM-WM	0.8052	**0.0061**
DM-DW	0.8367	**0.0152**
DM-MS	0.1935	**0.0118**
DW-WM	0.7701	**0.0020**
MS-WM	**< 0.0001**	0.4198
DW-MS	**0.0017**	**0.0015**

3. Normalized Absolute Difference (NAD)

a. CBD

Figure 3-25 compares NAD values for daytime and nighttime JU03 predictions within the CBD using the PNA MET input option and the five Urban HPAC modes that we examined. Table 3-17 presents the corresponding hypothesis test results, which are summarized below.

- For the day releases, the predictions associated with **UC**, **WM**, **DM**, and **DW** have smaller (less scatter) NAD values than those associated with **MS**. Also, the **UC**, **WM**, and **DW** predictions exhibited less scatter (as measured by NAD) than the **DM**-mode predictions.

- At night, the picture is mixed with several small differences being detectable as follows: **DM** and **DW** improved relative to **MS,** and **MS** improved relative to **UC** and **WM**.

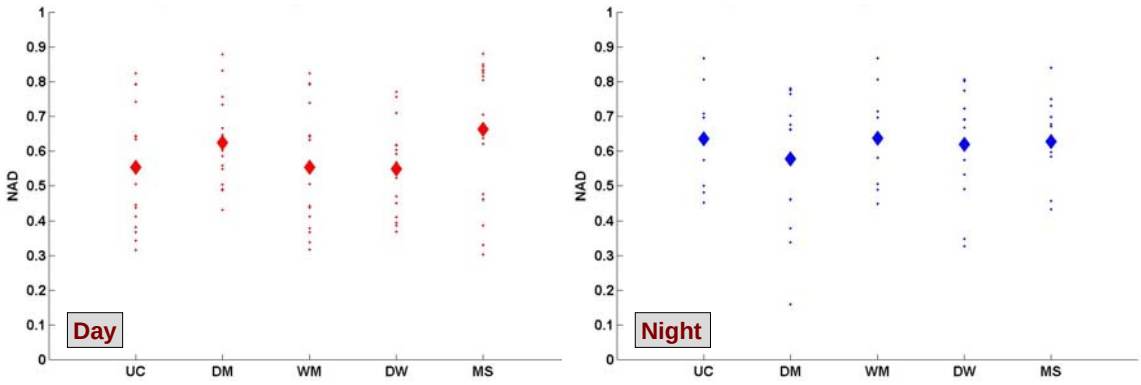


Figure 3-25. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations within the CBD using the PNA MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-17. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the PNA MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	**0.0103**	0.3816
WM-UC	0.1963	0.1670
DW-UC	0.8811	0.7265
MS-UC	**0.0001**	**0.0044**
DM-WM	**0.0084**	0.3666
DM-DW	**0.0001**	0.0611
DM-MS	**0.0003**	**0.0063**
DW-WM	0.9030	0.6920
MS-WM	**0.0002**	**0.0043**
DW-MS	**< 0.0001**	**0.0090**

b. Arcs

Figure 3-26 compares NAD values for daytime and nighttime JU03 predictions on the arcs using the PNA MET input option. Hypothesis test results for NAD comparisons (on the arcs) are shown in Table 3-18. The only significant finding was for the daytime **DW**-**DM** comparison, with the **DW** mode resulting in less (albeit a very small difference) scatter than the **DM** mode.

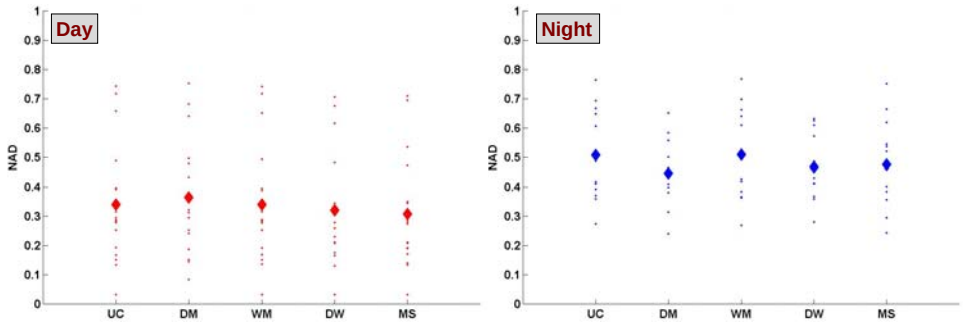


Figure 3-26. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations on the arcs using the PNA MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-18. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the PNA MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.1998	0.1665
WM-UC	0.3651	0.5871
DW-UC	0.0916	0.3245
MS-UC	0.1285	0.1448
DM-WM	0.1998	0.1515
DM-DW	**0.0105**	0.2251
DM-MS	0.0690	0.4629
DW-WM	0.0971	0.3143
MS-WM	0.1201	0.1306
DW-MS	0.7778	0.9719

4. Normalized Mean Square Error (NMSE)

a. CBD

Figure 3-27 compares NMSE values for daytime and nighttime JU03 predictions within the CBD using the PNA MET input option. Table 3-19 presents the hypothesis test results that compare NMSE values for daytime and nighttime JU03 predictions within the CBD using the PNA MET input option. For the daytime releases and predictions, **UC, WM**, **DM**, and **DW** result in lower scatter than **MS**, and the **DW** mode led to improved scatter (as measured by NMSE) relative to the **DM** mode. For the night predictions, we found that no differences (among the 10 possible comparisons) could be discerned (at the 0.05 significance level).

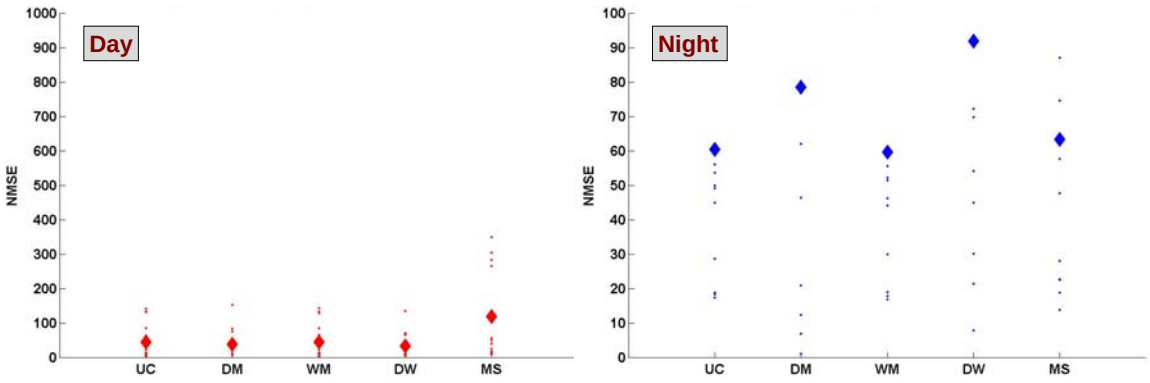


Figure 3-27. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations within the CBD using the PNA MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. Some of the larger values are off the depicted scale. The larger colored diamonds correspond to the average NMSE value.

b. Arcs

Figure 3-28 compares NMSE values for daytime and nighttime JU03 predictions on the arcs using the PNA MET input option. Hypothesis test results for NMSE comparisons (on the arcs) are shown in Table 3-20. These test results suggest:

- During the day, **MS** and **DW** result in lower scatter than **UC** and **WM**.
- At night, a slight improvement in **UC** relative to **WM** is detectable. While this difference is detectable with our paired hypothesis testing procedures, it is not

considered particularly important since the magnitude of the differences is so small.

Table 3-19. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the PNA MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.4955	0.2152
WM-UC	0.6680	0.3263
DW-UC	0.1543	0.0825
MS-UC	**0.0133**	0.2910
DM-WM	0.4748	0.1947
DM-DW	**0.0291**	0.0863
DM-MS	**0.0049**	0.7615
DW-WM	0.1528	0.0773
MS-WM	**0.0145**	0.2667
DW-MS	**0.0017**	0.9431

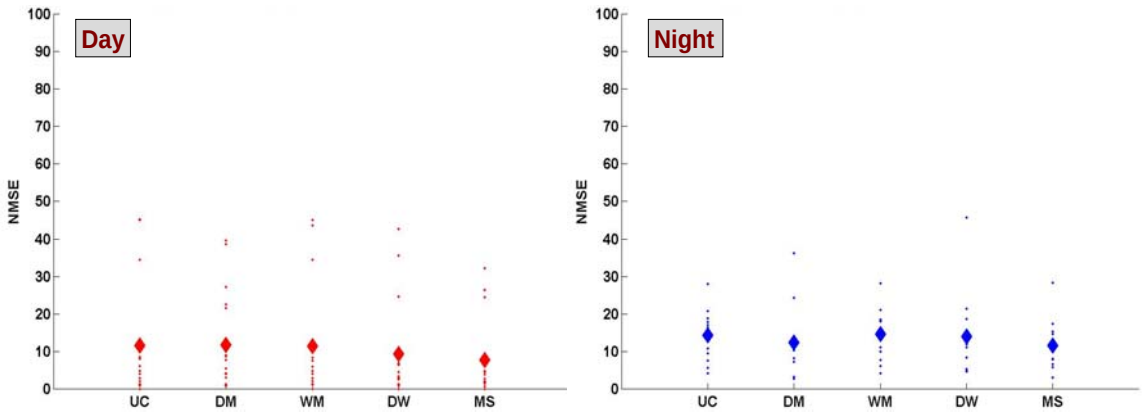


Figure 3-28. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations on the arcs using the PNA MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NMSE value.

Table 3-20. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the PNA MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.8537	0.3843
WM-UC	0.2611	**0.0356**
DW-UC	**0.0149**	0.9057
MS-UC	**0.0214**	0.2785
DM-WM	0.7928	0.3407
DM-DW	0.1017	0.2813
DM-MS	0.0883	0.9903
DW-WM	**0.0216**	0.8458
MS-WM	**0.0203**	0.2378
DW-MS	0.4749	0.4633

5. Threshold-Based MOE

a. CBD

Figure 3-29 compares threshold-based MOE values for daytime and nighttime JU03 predictions within the CBD using the PNA MET input option. Two thresholds are considered 25 and 250 ppt, which correspond to 5 and 50 × the estimated SF_6 background. For the day releases, all five Urban HPAC modes resulted in under-predictions at both thresholds, albeit quite minimal in the case of **DM** at 25 ppt and **MS** at 250 ppt. For the night releases, the Urban HPAC modes generally led to mild over-predictions (of the number of samplers that exceeded the assessed threshold), with the exceptions of **DM** and **DW** at 250 ppt, which predict the right amount of samplers (i.e., the confidence region clusters for the 250 ppt threshold MOE values for these modes straddle the diagonal in MOE space).

No statistically significant differences – in terms of improved false positive *and* improved false negative – could be discerned for these CBD-based (day or night) threshold-MOE comparisons. The overall improvement in MOE value for the threshold-based computations relative to the concentration-based values, as discussed for other MET input options, can be seen by comparing Figures 3-21 and 3-29.)

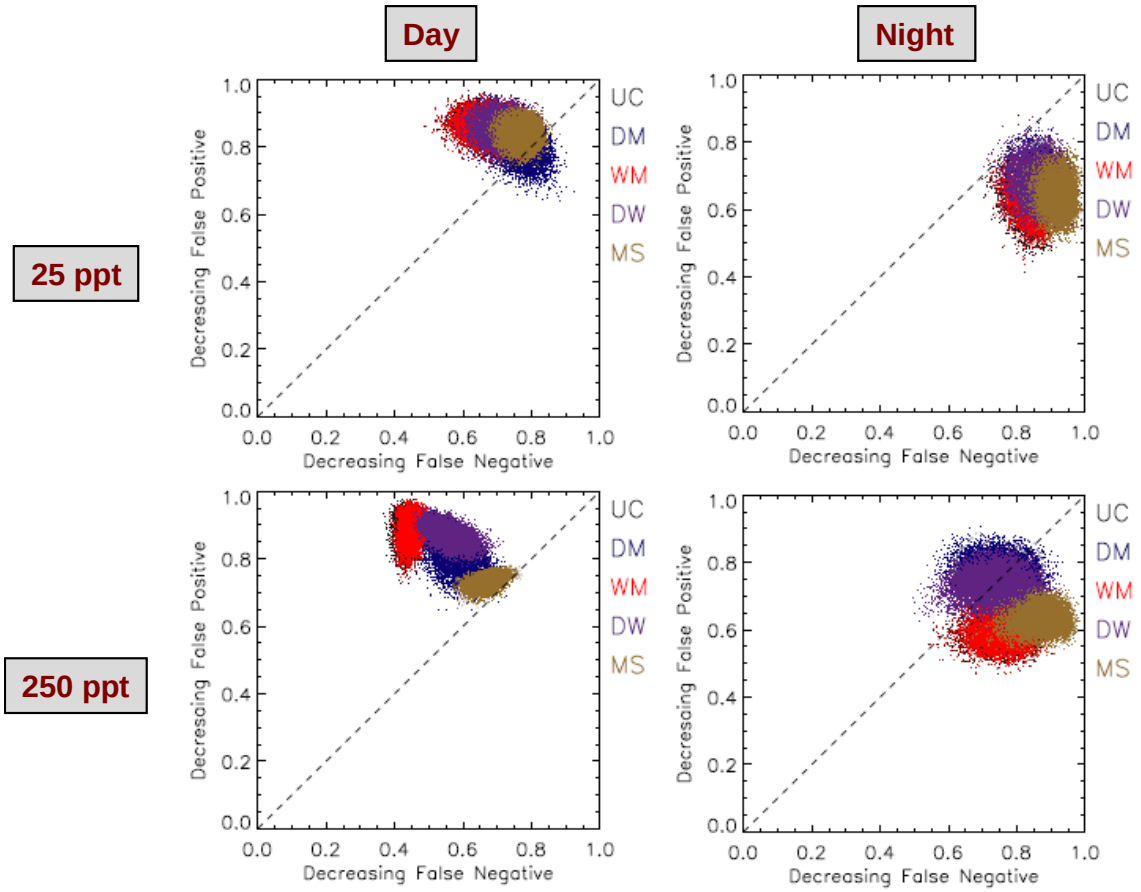


Figure 3-29. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the PNA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

b. Arcs

Figure 3-30 compares threshold-based MOE values for daytime and nighttime JU03 predictions along the arcs using the PNA MET input option. For the day releases, and the 25 ppt threshold, all five Urban HPAC modes led to MOE values that cluster near the diagonal. At 250 ppt, all five modes resulted in moderate under-predictions. At night, the Urban HPAC modes resulted in over-predictions for both thresholds with the exception of **DM** and **DW** at 250 ppt, which predict the right amount of samplers. These "on-the-arcs" results are similar to those found for the BAS and BRB MET options. With respect to hypothesis test results, no significant improvements were discerned during the day, and at night, **DM** was found to lead to an improved 25-ppt based threshold MOE relative to **UC**.

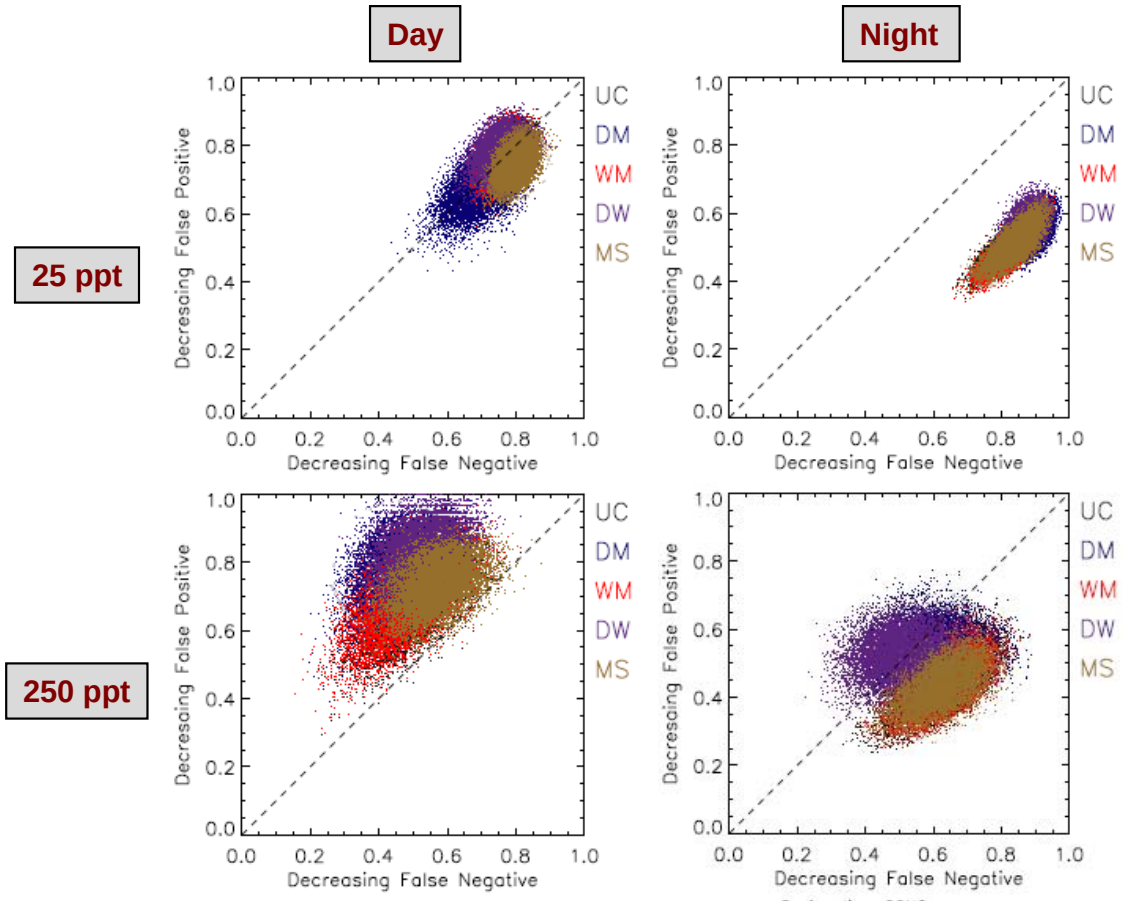


Figure 3-30. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the PNA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

6. Summary for PNA MET Input Option Results

Based on the above discussions and consideration of the comparisons of the five metrics, particularly the measures of scatter (NAD, NMSE, and concentration-based MOE), the qualitative results shown in Table 3-21 appear reasonable. In general, mixed results were found. However, during the day in the CBD, **UC**, **WM**, and **DW** offered less scatter relative to **MS**, albeit at the cost of a large under-prediction relative to **MS** (Figure 3-21). That is, even the statistically detectable difference in scatter metrics may be of questionable value in terms of assessing overall model performance because the bias difference is so large. In addition, during the day and in the CBD, **DW** outperformed **DM**. These PNA-based results should be thought of as consistent with an

Urban HPAC usage that employs a single observed upwind profile as meteorological input – a possibly operationally representative use of Urban HPAC for some specific scenarios.

Table 3-21. Urban HPAC Modes (PNA MET) That Consistently, Across at Least Two of the Three Scatter-Related Metrics, Led to Improved Predictive Performance of JU03 Releases

Surface Sampler Region	Day	Night
CBD	(UC,WM,DM,DW)/MS and DW/DM	mixed
Arcs	mixed	mixed

D. COMPARSIONS FOR THE ACA MET OPTION

1. Concentration-Based MOE

a. CBD

Figure 3-31 compares concentration-based MOE values for daytime and nighttime JU03 predictions within the CBD using the ACA MET input option. For the day releases, four of the Urban HPAC modes led to under-predictions. The exception is the **MS** mode, which resulted in an over-prediction. No significant differences were discerned within the CBD (day or night) using the previously described hypothesis test procedures. These ACA-based results are very similar to the PNA-based results (Figure 3-21).

b. Arcs

Figure 3-32 compares concentration-based MOE values for daytime and nighttime JU03 predictions along the arcs using the ACA MET input option. No significant differences were discerned for either the day or night releases when considering the samplers along the arcs. The ACA-based results shown here are very similar to the PNA-based results (Figure 3-22).

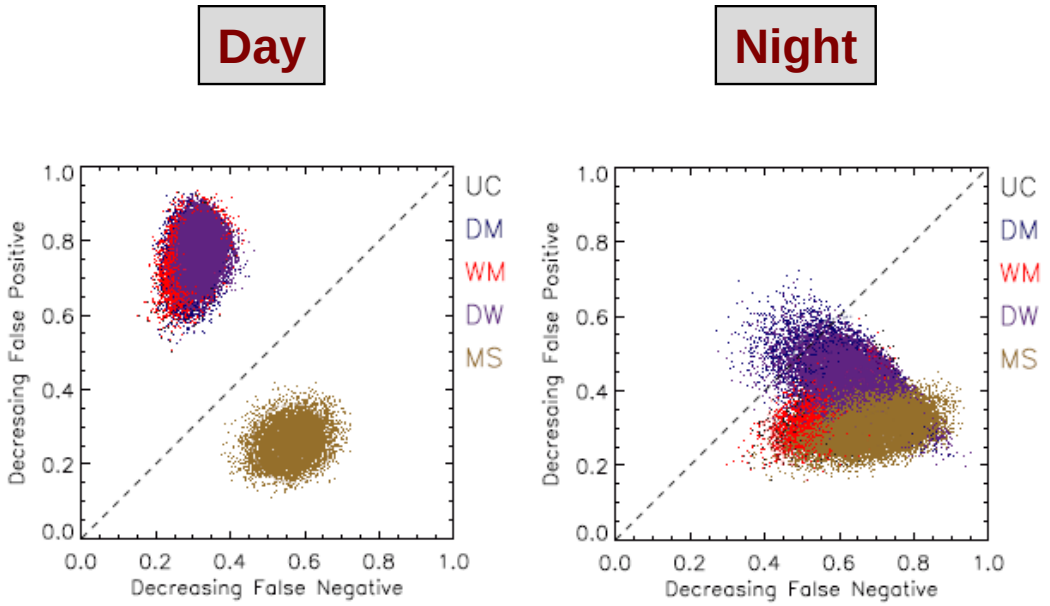


Figure 3-31. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the ACA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

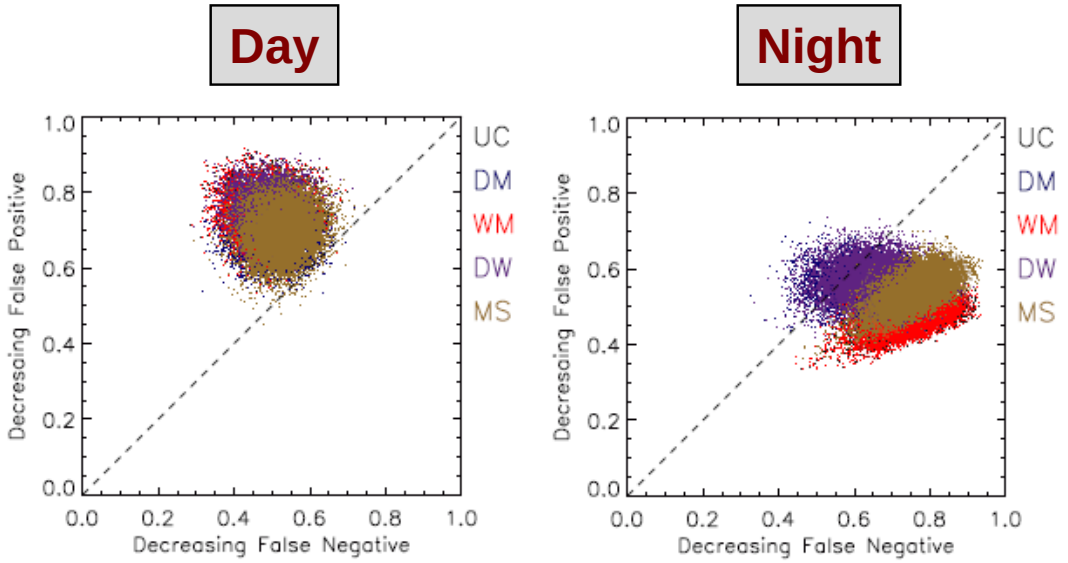


Figure 3-32. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the ACA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

2. Fractional Bias (FB)

a. CBD

Figure 3-33 compares FB values for daytime and nighttime JU03 predictions within the CBD using the ACA MET input option. As seen previously with the concentration-based MOE, for the day releases, four of the Urban HPAC modes led to under-predictions (FB < 0). The exception is the **MS** mode, which led to a significant over-prediction. This same result was seen for the PNA MET input option predictions in the CBD. At night, all five modes led to over-predictions. Hypothesis test results for FB comparisons (in the CBD) are shown in Table 3-22. The following significant findings are obtained:

- For the day releases, the **MS** FB values are found to be significantly different (over-predictions) from those associated with the **UC**, **DM**, **WM**, and **DW** predictions (all under-predictions). In addition, a very slight improvement in **UC** relative to **WM** is detectable. While this difference is detectable with our paired hypothesis testing procedures, it is not considered particularly important since the magnitude of the difference is so small.
- For the nighttime releases and predictions, no statistically significant differences could be discerned.

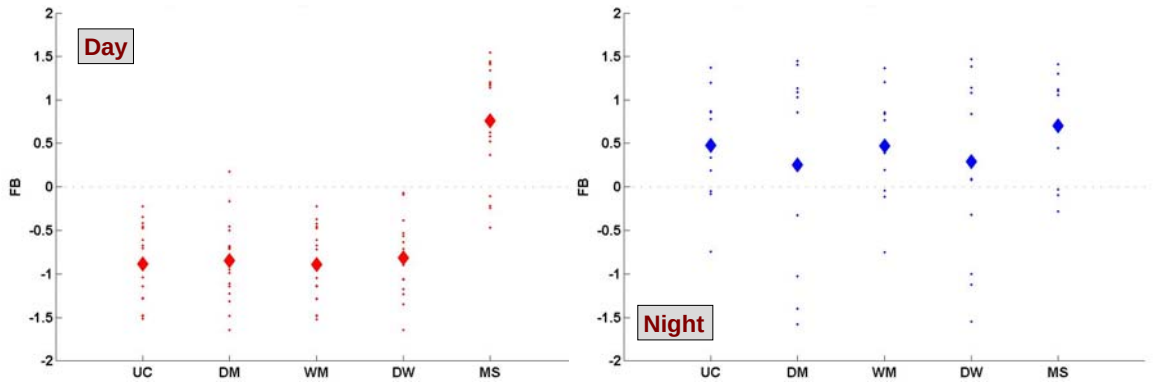


Figure 3-33. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations within the CBD using the ACA MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

Table 3-22. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the ACA MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.7694	0.2379
WM-UC	**0.0039**	0.6282
DW-UC	0.5929	0.2902
MS-UC	**< 0.0001**	0.2613
DM-WM	0.7580	0.2452
DM-DW	0.5410	0.3296
DM-MS	**< 0.0001**	0.1669
DW-WM	0.5712	0.2966
MS-WM	**< 0.0001**	0.2517
DW-MS	**< 0.0001**	0.2070

b. Arcs

Figure 3-34 compares FB values for daytime and nighttime JU03 predictions on the arcs using the ACA MET input option. Hypothesis test results for FB comparisons on the arcs are shown in Table 3-23 and suggest the following:

- For the day releases, the **MS** FB values are found to be significantly different (and improved, i.e., less under-prediction) from those associated with the **UC**, **WM**, and **DW** predictions.
- At night, the **DM** and **DW** modes resulted in less over-prediction relative to **UC** and **WM**. Also, the FB values associated with the **DW** mode predictions are found to be different from those of **MS**, with **DW** resulting in less over-prediction.

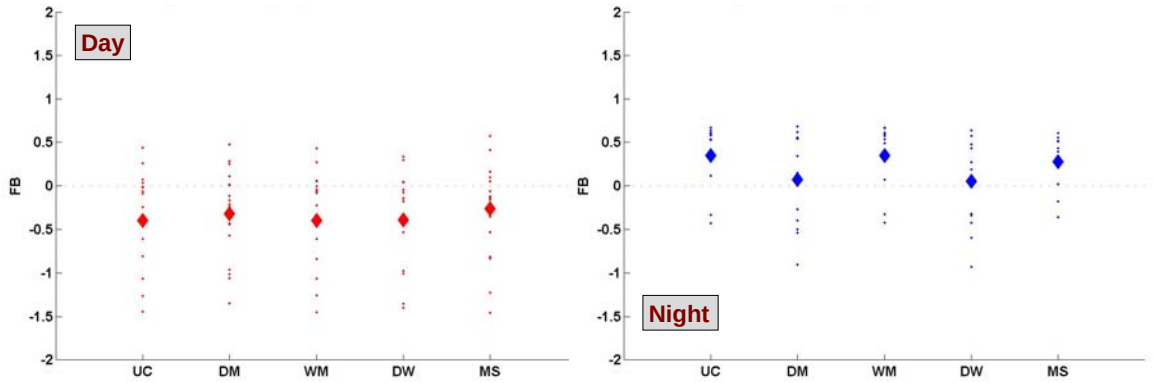


Figure 3-34. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations on the arcs using the ACA MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

Table 3-23. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the ACA MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.2955	**0.0276**
WM-UC	0.8521	0.4800
DW-UC	0.8891	**0.0026**
MS-UC	**0.0001**	0.0763
DM-WM	0.3065	**0.0271**
DM-DW	0.1438	0.7448
DM-MS	0.4128	0.0832
DW-WM	0.8799	**0.0027**
MS-WM	**0.0002**	0.0825
DW-MS	**0.0016**	**0.0343**

3. Normalized Absolute Difference (NAD)

a. CBD

Figure 3-35 compares NAD values for daytime and nighttime JU03 predictions within the CBD using the ACA MET input option and the five Urban HPAC modes that we examined. Table 3-24 presents the corresponding hypothesis test results, which are summarized below.

- For the day releases, the predictions associated with **DW** resulted in smaller (less scatter) NAD values than those associated with **MS**.
- For the nighttime releases and predictions, no statistically significant differences could be discerned.

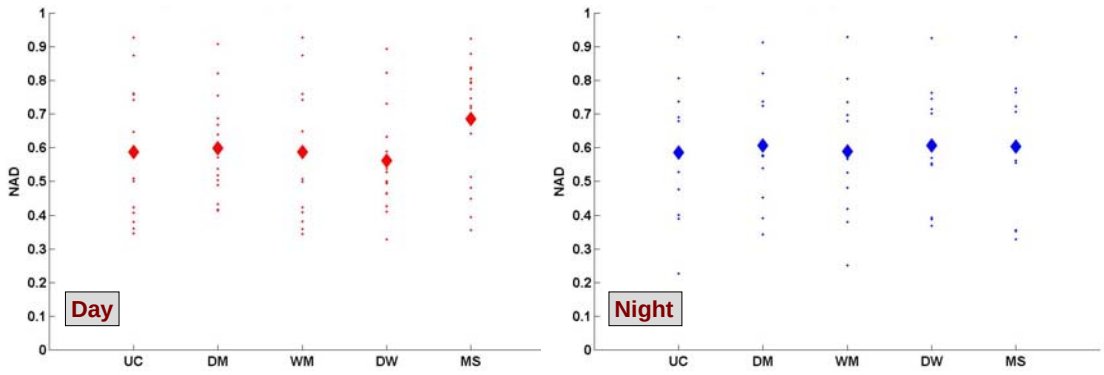


Figure 3-35. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations within the CBD using the ACA MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

b. Arcs

Figure 3-36 compares NAD values for daytime and nighttime JU03 predictions on the arcs using the ACA MET input option. Hypothesis test results for NAD comparisons (on the arcs) are shown in Table 3-25 and are summarized below.

- For the daytime releases and predictions, no statistically significant differences could be discerned.
- At night, the predictions associated with **MS** resulted in less scatter than those associated with **UC**, **WM**, and **DM**.

Table 3-24. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the ACA MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.7712	0.5011
WM-UC	0.7671	0.3717
DW-UC	0.4449	0.5245
MS-UC	0.1719	0.8019
DM-WM	0.7608	0.5277
DM-DW	0.1790	0.9841
DM-MS	0.1102	0.9763
DW-WM	0.4418	0.5806
MS-WM	0.1676	0.8239
DW-MS	**0.0245**	0.9704

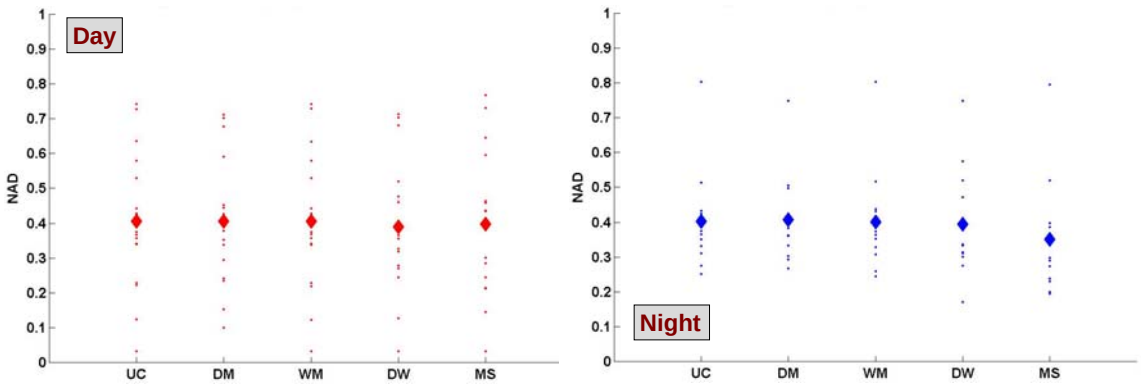


Figure 3-36. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations on the arcs using the ACA MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-25. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the ACA MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.9850	0.8210
WM-UC	0.7679	0.5199
DW-UC	0.2326	0.7469
MS-UC	0.5704	**0.0063**
DM-WM	0.9622	0.7735
DM-DW	0.1107	0.4726
DM-MS	0.6721	**0.0328**
DW-WM	0.2459	0.7759
MS-WM	0.5648	**0.0040**
DW-MS	0.6144	0.0767

4. Normalized Mean Square Error (NMSE)

a. CBD

Figure 3-37 compares NMSE values for daytime and nighttime JU03 predictions within the CBD using the ACA MET input option. Table 3-26 presents the corresponding hypothesis test results, which are summarized below.

- For the day releases, the predictions associated with **DW** resulted in smaller (less scatter) NMSE values than those associated with **MS**. This conclusion is similar to the corresponding NAD finding.
- For the nighttime releases and predictions, no statistically significant differences could be discerned.

b. Arcs

Figure 3-38 compares NMSE values for daytime and nighttime JU03 predictions on the arcs using the ACA MET input option. Hypothesis test results for NMSE comparisons (on the arcs) are shown in Table 3-27. These test results suggest:

- During the day, no statistically significant differences could be discerned.

- At night, the predictions associated with **MS** resulted in less scatter than those associated with **UC**, **WM**, and **DM**.

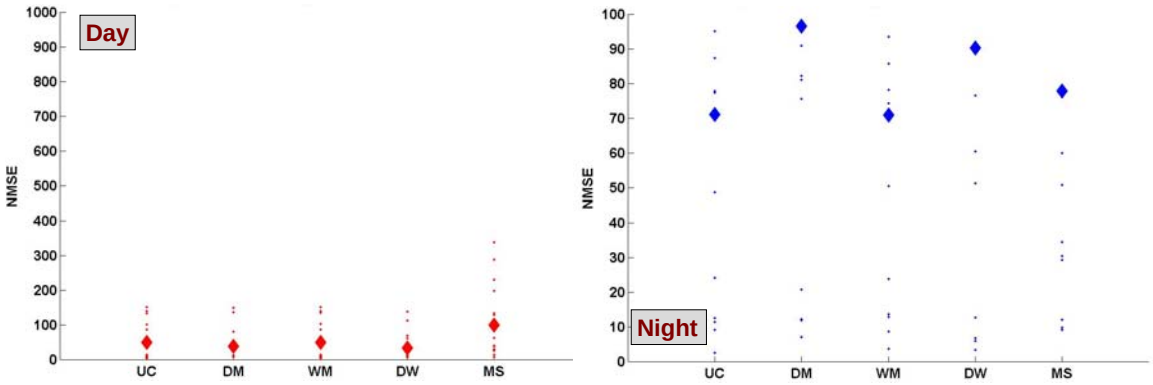


Figure 3-37. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations within the CBD using the ACA MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. Some of the larger values are off the depicted scale. The larger colored diamonds correspond to the average NMSE value.

Table 3-26. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the ACA MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.3601	0.1203
WM-UC	0.2216	0.9291
DW-UC	0.1741	0.2481
MS-UC	0.1780	0.8652
DM-WM	0.3645	0.0963
DM-DW	0.1389	0.2812
DM-MS	0.0645	0.6675
DW-WM	0.1734	0.2402
MS-WM	0.1728	0.8529
DW-MS	**0.0448**	0.7768

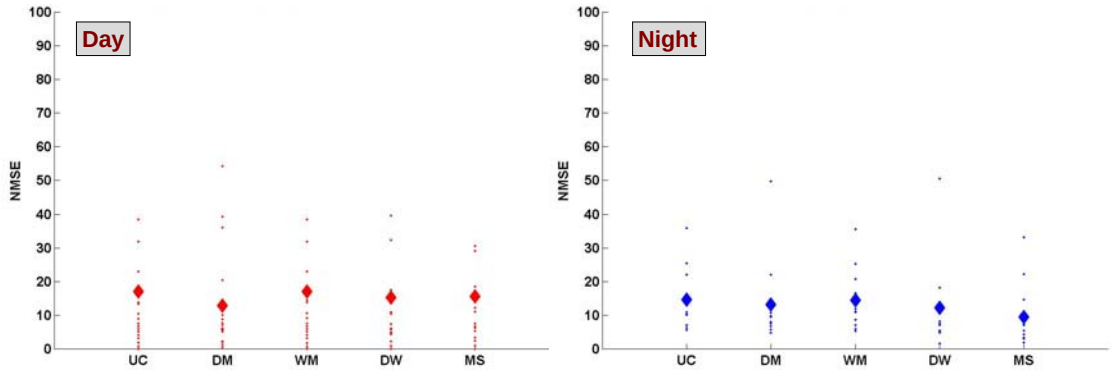


Figure 3-38. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations on the arcs using the ACA MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NMSE value.

Table 3-27. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the ACA MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.2669	0.4696
WM-UC	0.1584	0.2821
DW-UC	0.1329	0.2751
MS-UC	0.1806	**0.0005**
DM-WM	0.2534	0.5423
DM-DW	0.8327	0.2897
DM-MS	0.8012	**0.0279**
DW-WM	0.1307	0.3409
MS-WM	0.1589	**0.0010**
DW-MS	0.8377	0.1176

5. Threshold-Based MOE

a. CBD

Figure 3-39 compares threshold-based MOE values for daytime and nighttime JU03 predictions within the CBD using the ACA MET input option. For the day

releases, the Urban HPAC modes generally resulted in under-predictions at both thresholds, albeit quite minimally in the case of the 25 ppt threshold and for **MS** at the 250 ppt threshold. For the night releases, the Urban HPAC modes led to relatively unbiased predictions, in terms of the number of samplers that exceeded the specified threshold. No statistically significant threshold-MOE improvements could be discerned for these CBD-based (day or night) comparisons. The overall improvement in MOE value for the threshold-based computations relative to the concentration-based values, as discussed for other MET input options, can be seen by comparing Figures 3-31 and 3-39. Finally, the ACA-based results shown here are similar to those associated with the PNA MET input option (Figure 3-29).

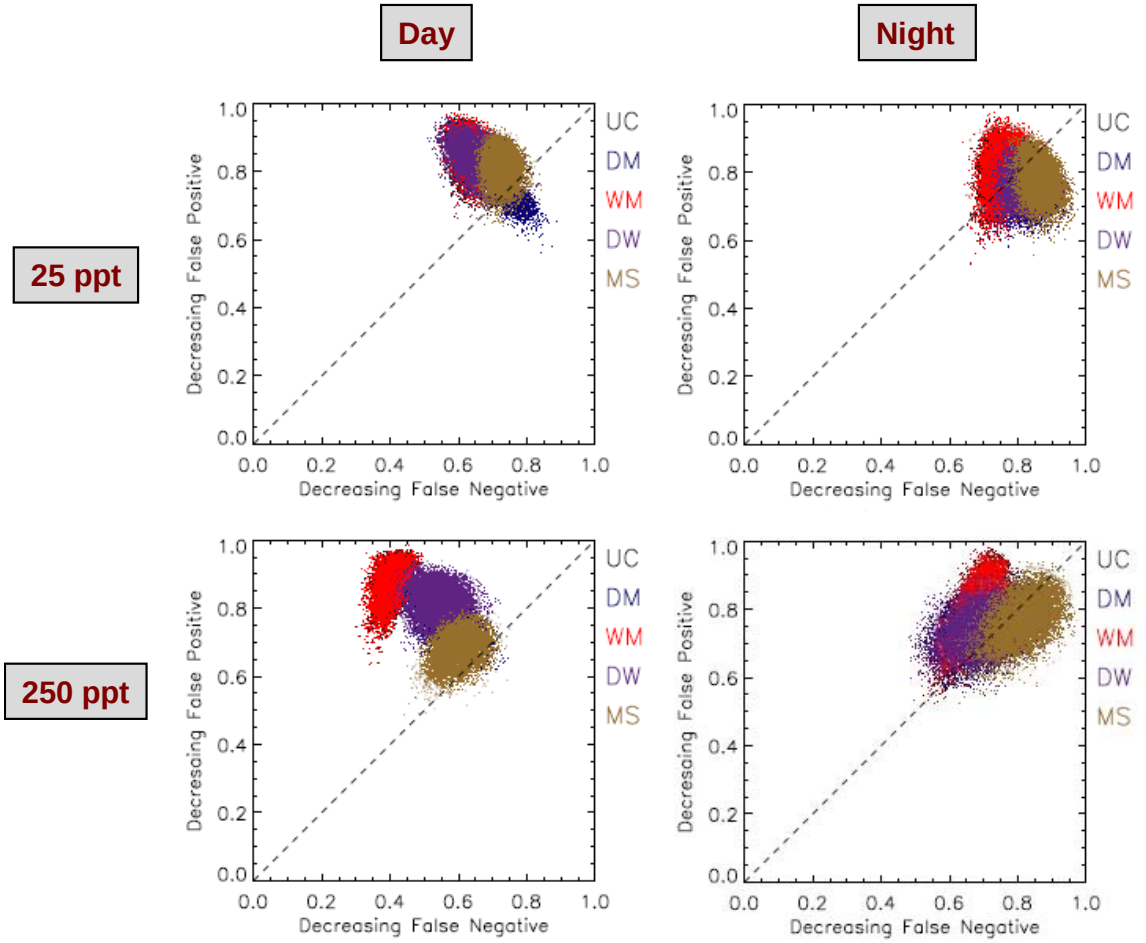


Figure 3-39. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the ACA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

b. Arcs

Figure 3-40 compares threshold-based MOE values for daytime and nighttime JU03 predictions along the arcs using the ACA MET input option. For the day releases, and the 25 ppt threshold, all five Urban HPAC modes led to MOE values that cluster near the diagonal. At 250 ppt and during the day, all five modes resulted in slight under-predictions. At night, the Urban HPAC modes generally resulted in over-predictions for both thresholds, with the exceptions being **DM** and **DW**, which appear relatively unbiased. No statistically significant threshold-MOE improvements could be discerned for these arc-based (day or night) comparisons. The ACA-based results shown here for the arc-base samplers are almost identical for those associated with the PNA MET input option (Figure 3-30).

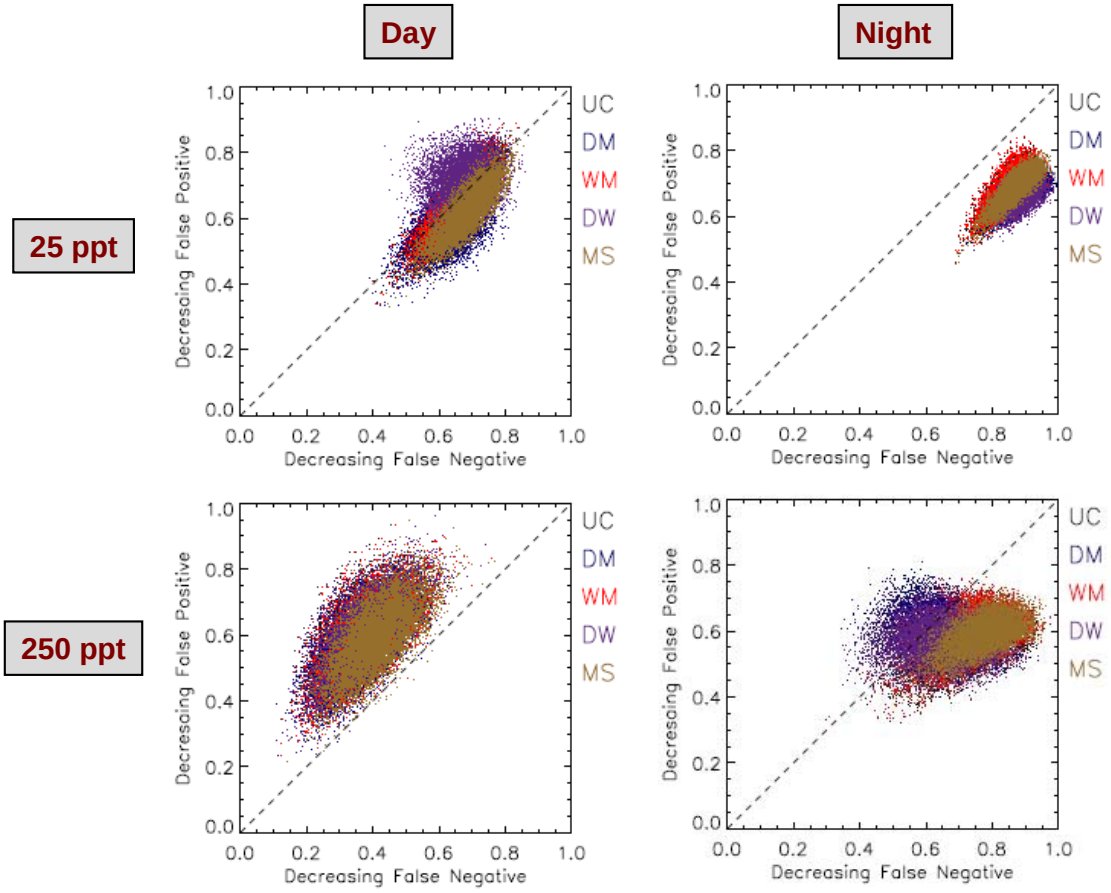


Figure 3-40. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the ACA MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

6. Summary for ACA MET Input Option Results

Based on the above discussions and consideration of the comparisons of the five metrics, particularly the measures of scatter (NAD, NMSE, and concentration-based MOE), the qualitative results shown in Table 3-28 appear reasonable. In general, mixed results were found. However, at night on the arcs (at longer range), the **MS** mode resulted in less scatter (as measured by NAD and NMSE) than **UC**, **WM**, and **DM**. Also, during the day in the CBD, **DW** outperformed **MS**. These ACA-based results should be thought of as consistent with an Urban HPAC usage that employs a single observed downwind profile as meteorological input – a possibly operationally representative use of Urban HPAC for some specific scenarios.

Table 3-28. Urban HPAC Modes (SCA MET) That Consistently, Across at Least Two of the Three Scatter-Related Metrics, Led to Improved Predictive Performance of JU03 Releases

Surface Sampler Region	Day	Night
CBD	**DW**/**MS**	no differences
Arcs	no differences	**MS** / (**UC**,**WM**,**DM**)

E. COMPARSIONS FOR THE PO7 MET OPTION

1. Concentration-Based MOE

a. CBD

Figure 3-41 compares concentration-based MOE values for daytime and nighttime JU03 predictions within the CBD using the PO7 MET input option and the five Urban HPAC modes that we examined. For the day releases, four of the Urban HPAC modes led to under-predictions ("above the diagonal"). The exception is the **MS** mode, which predicted about the right amount of material at the surface samplers. Using the hypothesis test procedure described in the previous chapter, we found that no differences (among the 10 possible comparisons) could be discerned (at the 0.0496 significance level as described previously for the 17 independent daytime releases).

At night, four of the five modes led to over-predictions with the **MS** mode being the partial exception (showing only a slight over-prediction). In terms of having an MOE value closest to the perfect value of (1,1), the nighttime results show that **MS**, **DW**, and **DM** are closer than **UC** and **WM**. Hypothesis testing revealed significant differences (at

the 0.0112 significance level), with the first mode listed being the one with values closer to (1,1) and having less false negative *and* less false positive, for the following six comparisons: **MS**-**UC**, **DM**-**UC**, **DW**-**UC**, **MS**-**WM**, **DM**-**WM**, and **DW**-**WM**.

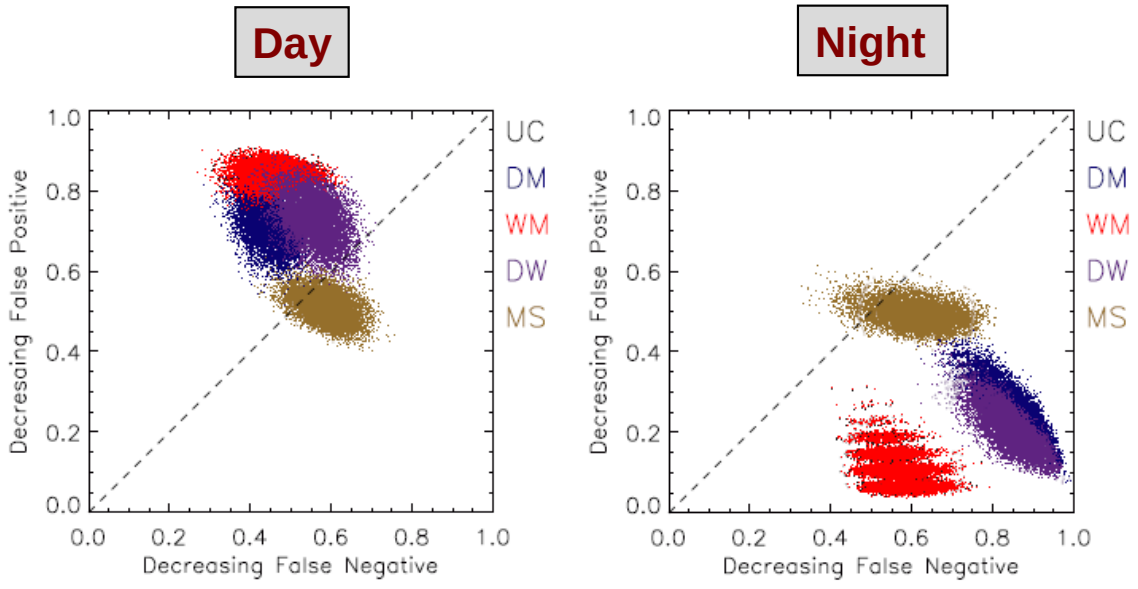


Figure 3-41. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the PO7 MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

b. Arcs

Figure 3-42 compares concentration-based MOE values for daytime and nighttime JU03 predictions along the arcs using the PO7 MET input option. For the day releases, all five Urban HPAC modes led to under-predictions, albeit slight in some cases. For the night releases, all five Urban HPAC modes led to over-predictions with the **MS** predictions being the smallest over-prediction. Hypothesis testing revealed slight (but detectable) improvements in the MOE values associated with daytime **DW** predictions on the arcs relative to the **UC** and **WM** predictions. For the night predictions, no significant improvements ("less false negative and less false positive") were suggested by the hypothesis testing using the concentration-based MOE and considering the samplers on the arcs.

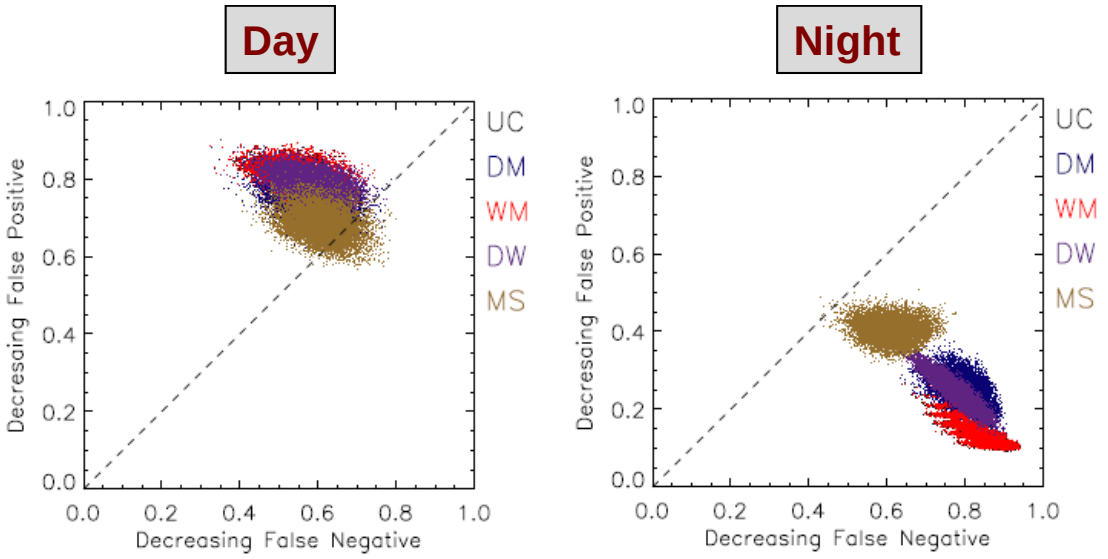


Figure 3-42. Comparisons of Concentration-Based MOE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the PO7 MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

2. Fractional Bias (FB)

a. CBD

Figure 3-43 compares FB values for daytime and nighttime JU03 predictions within the CBD using the PO7 MET input option. As seen previously with the concentration-based MOE, for the day releases, four of the Urban HPAC modes led to under-predictions (FB < 0). The exception is the **MS** mode, which led to a relatively unbiased prediction. At night, four of the modes led to over-predictions with the **MS** mode being the exception (showing relatively unbiased performance). Hypothesis test results for FB comparisons (in the CBD) are shown in Table 3-29 and summarized below.

- During the day, the **MS** and **DW** modes result in less under-prediction than the **UC**, **WM**, and **DM** modes. In addition, **MS** led to less under-prediction than **DW**.
- At night, the predictions associated with the **MS**, **DW**, and **DM** modes led to less over-prediction than the **UC** and **WM** mode predictions. In addition, **MS** resulted in less over-prediction than **DM** and **DW**.

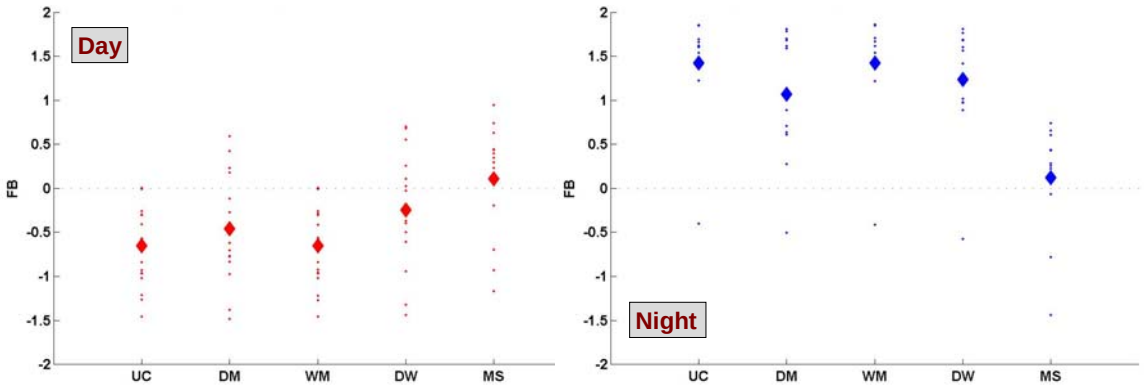


Figure 3-43. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations within the CBD using the PO7 MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

Table 3-29. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the PO7 MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.2919	**0.0346**
WM-UC	0.1894	0.8375
DW-UC	**0.0152**	**0.0216**
MS-UC	**< 0.0001**	**0.0004**
DM-WM	0.3015	**0.0337**
DM-DW	**0.0010**	0.0645
DM-MS	**0.0059**	**0.0011**
DW-WM	**0.0145**	**0.0197**
MS-WM	**< 0.0001**	**0.0002**
DW-MS	**0.0489**	**0.0007**

b. Arcs

Figure 3-44 compares FB values for daytime and nighttime JU03 predictions on the arcs using the PO7 MET input option. Hypothesis test results for FB comparisons on the arcs are shown in Table 3-30 and are summarized below.

- During the day, the **MS**, **DW**, and **DM** modes result in less under-prediction than the **UC** and **WM** modes.
- At night, the predictions associated with the **MS**, **DW**, and **DM** modes led to less over-prediction than the **UC** and **WM** mode predictions. In addition, **MS** resulted in less over-prediction than **DM** and **DW**.

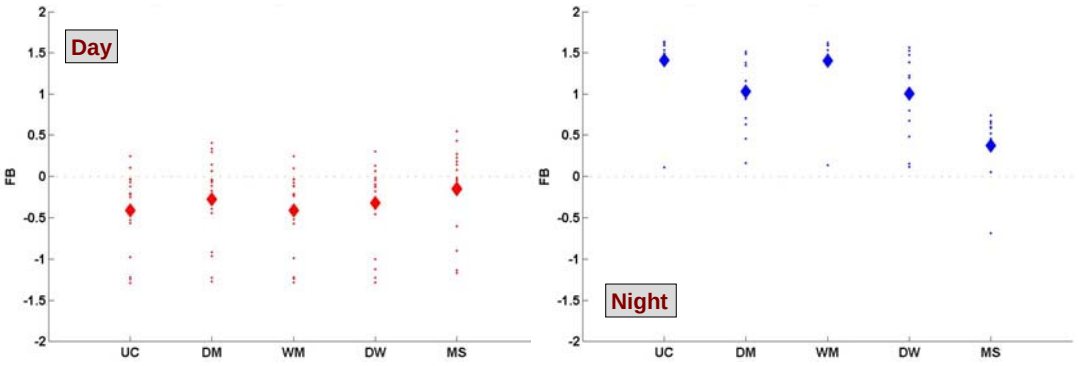


Figure 3-44. Comparisons of FB Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These FB values are for predictions of 30-minute average concentrations on the arcs using the PO7 MET input option.

The smaller colored (red for day and blue for night) points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB value.

Table 3-30. P-Values for Urban HPAC Mode Comparisons of FB for Day and Night Predictions of JU03 Releases. For these comparisons, the PO7 MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	**< 0.0001**	**0.0011**
WM-UC	0.4959	0.2612
DW-UC	**0.0437**	**0.0009**
MS-UC	**0.0004**	**0.0005**
DM-WM	**< 0.0001**	**0.0016**
DM-DW	0.1854	0.7442
DM-MS	0.2023	**0.0002**
DW-WM	**0.0370**	**0.0011**
MS-WM	**0.0001**	**0.0007**
DW-MS	0.0616	**0.0005**

3. Normalized Absolute Difference (NAD)

a. CBD

Figure 3-45 compares NAD values for daytime and nighttime JU03 predictions within the CBD using the PO7 MET input option. Table 3-31 presents the corresponding hypothesis test results, which are summarized below.

- During the day, the **UC**, **WM**, and **DW** modes resulted in less scatter than the **DM** mode.
- At night, the predictions associated with the **MS**, **DW**, and **DM** modes led to less scatter than the **UC** and **WM** mode predictions. In addition, **MS** resulted in less scatter than **DM** and **DW**, and **DM** led to less scatter than **DW**.

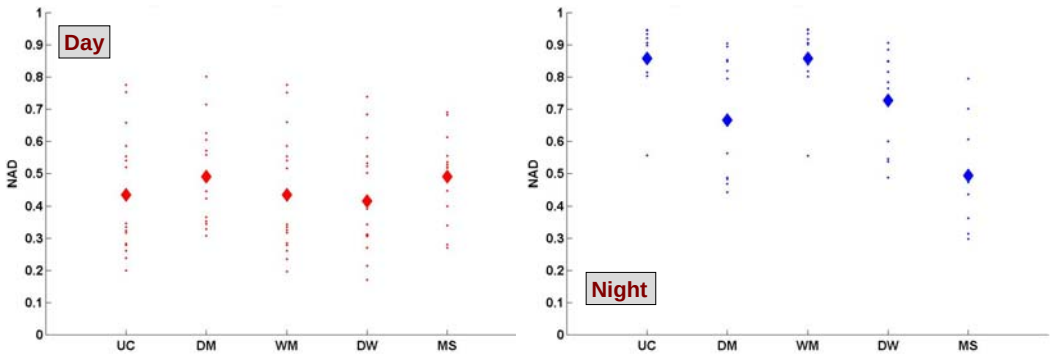


Figure 3-45. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations within the CBD using the PO7 MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

b. Arcs

Figure 3-46 compares NAD values for daytime and nighttime JU03 predictions on the arcs using the PO7 MET input option. Hypothesis test results for NAD comparisons (on the arcs) are shown in Table 3-32. These test results suggest:

- During the day, no statistically significant differences could be discerned.
- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**. In addition, **MS** represented an improvement (less scatter) relative to **DM** and **DW**. Finally, a very slight improvement in **WM** relative to **UC** is detectable. While

this difference is detectable, it is not considered particularly important since the magnitude of the differences is so small.

Table 3-31. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the PO7 MET input option was used and the CBD was considered

Comparison Tested	Day	Night
DM-UC	**0.0254**	**0.0008**
WM-UC	0.3208	0.9417
DW-UC	0.2516	**0.0007**
MS-UC	0.2297	**0.0017**
DM-WM	**0.0211**	**0.0003**
DM-DW	**0.0019**	**0.0339**
DM-MS	0.9803	**0.0732**
DW-WM	0.2624	**0.0004**
MS-WM	0.2191	**0.0018**
DW-MS	0.1150	**0.0113**

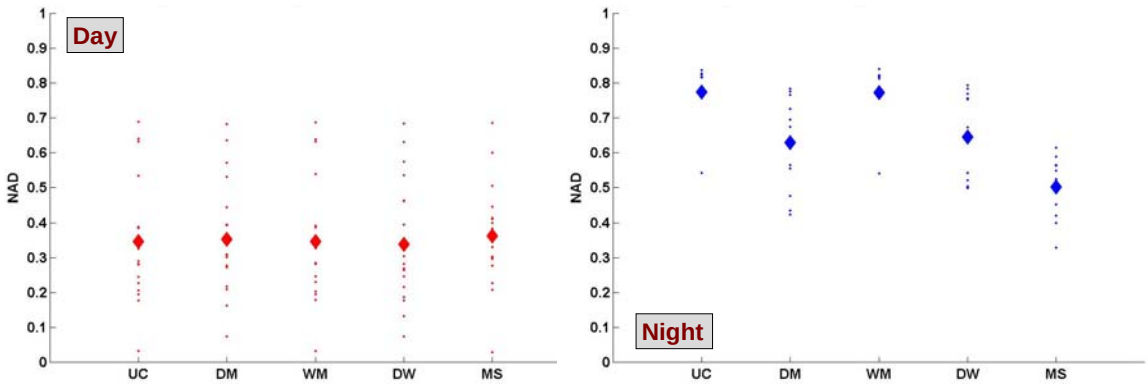


Figure 3-46. Comparisons of NAD Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NAD values are for predictions of 30-minute average concentrations on the arcs using the PO7 MET input option.

The smaller colored (red for day and blue for night) points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD value.

Table 3-32. P-Values for Urban HPAC Mode Comparisons of NAD for Day and Night Predictions of JU03 Releases. For these comparisons, the PO7 MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	0.6036	**0.0008**
WM-UC	0.6817	**0.0282**
DW-UC	0.4555	**0.0006**
MS-UC	0.3785	**0.0005**
DM-WM	0.6058	**0.0011**
DM-DW	0.0531	0.4204
DM-MS	0.6679	**0.0149**
DW-WM	0.4297	**0.0001**
MS-WM	0.3935	**0.0010**
DW-MS	0.2790	**0.0039**

4. Normalized Mean Square Error (NMSE)

a. CBD

Figure 3-47 compares NMSE values for daytime and nighttime JU03 predictions within the CBD using the PO7 MET input option. Table 3-33 presents the hypothesis test results that compare NMSE values for daytime and nighttime JU03 predictions within the CBD using the PO7 MET input option. These test results suggest:

- During the day, no statistically significant differences could be discerned.
- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**. In addition, **MS** represented an improvement (less scatter) relative to **DM** and **DW**.

b. Arcs

Figure 3-48 compares NMSE values for daytime and nighttime JU03 predictions on the arcs using the PO7 MET input option. Hypothesis test results for NMSE comparisons (on the arcs) are shown in Table 3-34. These test results suggest:

- For the daytime releases and predictions, the **DM** mode resulted in less scatter (as measured by NMSE) than the **UC** and **WM** modes.

- At night, **MS**, **DM**, and **DW** result in lower scatter than **UC** and **WM**. In addition, **MS** represented an improvement (less scatter) relative to **DM** and **DW**.

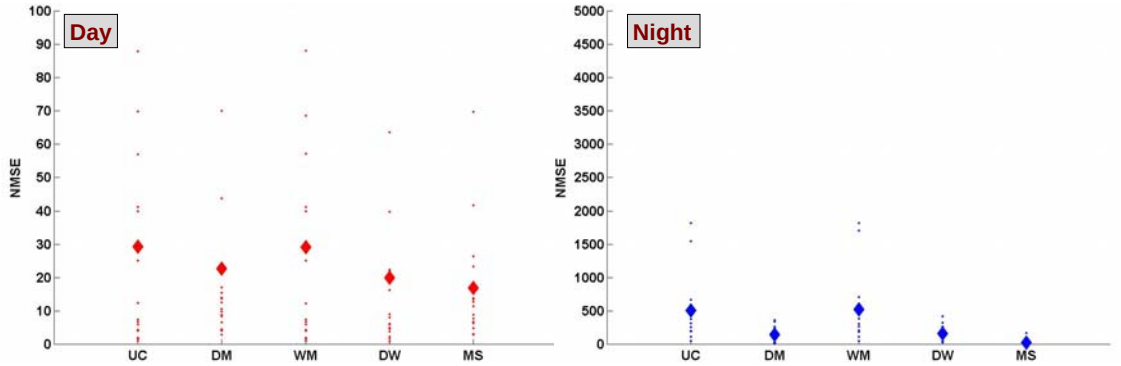


Figure 3-47. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations within the CBD using the PO7 MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. Some of the larger values are off of the depicted scale. The larger colored diamonds correspond to the average NMSE value.

Table 3-33. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the PO7 MET input option was used and the CBD was considered.

Comparison Tested	Day	Night
DM-UC	0.3088	**0.0044**
WM-UC	0.7349	0.3270
DW-UC	0.0560	**0.0103**
MS-UC	0.0769	**0.0014**
DM-WM	0.2971	**0.0053**
DM-DW	0.0854	0.1537
DM-MS	0.3319	**0.0309**
DW-WM	0.0588	**0.0093**
MS-WM	0.0724	**0.0017**
DW-MS	0.5445	**0.0205**

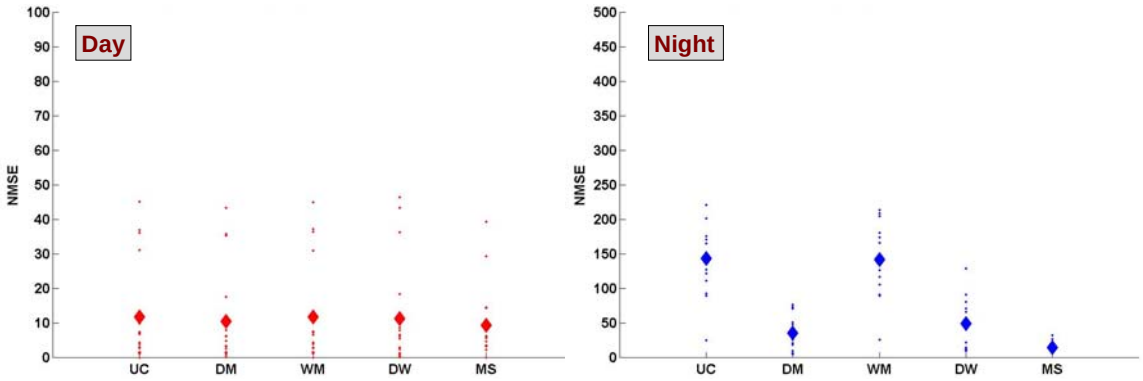


Figure 3-48. Comparisons of NMSE Values for Urban HPAC Predictions of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These NMSE values are for predictions of 30-minute average concentrations on the arcs using the PO7 MET input option.

The smaller colored (red for day and blue for night) points correspond to NMSE values for each of the individual releases (17 day and 12 night). Note the value scales are different for day and night. The larger colored diamonds correspond to the average NMSE value.

Table 3-34. P-Values for Urban HPAC Mode Comparisons of NMSE for Day and Night Predictions of JU03 Releases. For these comparisons, the PO7 MET input option was used and the arcs were considered.

Comparison Tested	Day	Night
DM-UC	**0.0238**	**0.0003**
WM-UC	0.7811	0.3687
DW-UC	0.5690	**0.0009**
MS-UC	0.2701	**0.0010**
DM-WM	**0.0225**	**0.0002**
DM-DW	0.1593	0.3036
DM-MS	0.6715	**0.0227**
DW-WM	0.5700	**0.0003**
MS-WM	0.2727	**0.0009**
DW-MS	0.5933	**0.0103**

5. Threshold-Based MOE

a. CBD

Figure 3-49 compares threshold-based MOE values for daytime and nighttime JU03 predictions within the CBD using the PO7 MET input option and the five Urban

HPAC modes that were examined. For the day releases and the 25 ppt threshold, all five Urban HPAC modes led to over-predictions, albeit quite mild in some cases. At 250 ppt and during the day, two modes (**UC**[8] and **WM**) resulted in under-predictions, one mode (**MS**) leads to a slight over-prediction, and two modes (**DM** and **DW**) straddle the diagonal (and hence are unbiased). For the nighttime predictions, the **MS**, **DM**, and **DW** modes led to over-predictions, albeit very slight for **MS** at 250 ppt, and the **UC** and **WM** modes resulted in under-predictions for both thresholds examined in Figure 3-49. We found no significant differences using the previously described four quadrant hypothesis test procedure for the MOE.

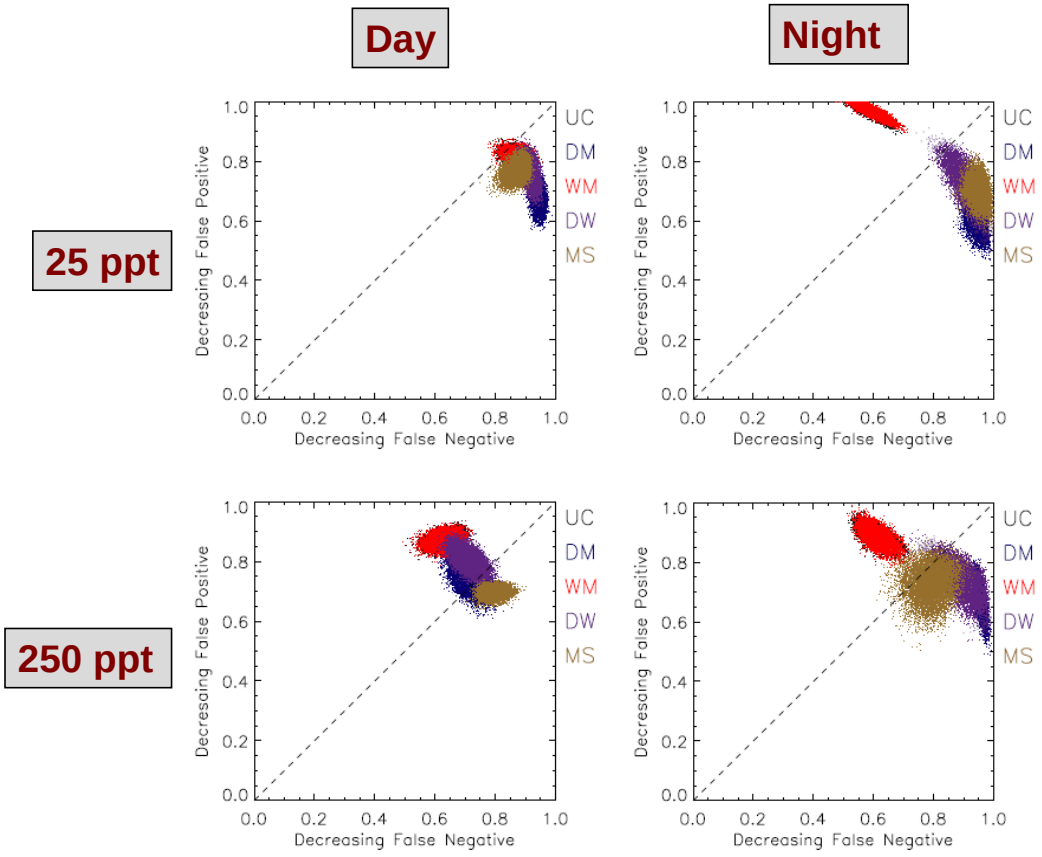


Figure 3-49. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations within the CBD using the PO7 MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

---

8 In Figure 3-49, the **UC** MOE confidence region (black cluster) is mainly obscured by the **WM** confidence region (red cluster) as they lie roughly on top of one another for both the day and night conditions.

b. Arcs

Figure 3-50 compares threshold-based MOE values for daytime and nighttime JU03 predictions along the arcs using the PO7 MET input option. During the day, all five Urban HPAC modes led to slight over-predictions at 25 ppt and under-predictions at 250 ppt. At night, all five Urban HPAC modes resulted in over-predictions for both thresholds. We found no significant differences using the previously described four quadrant hypothesis test procedure for the MOE.

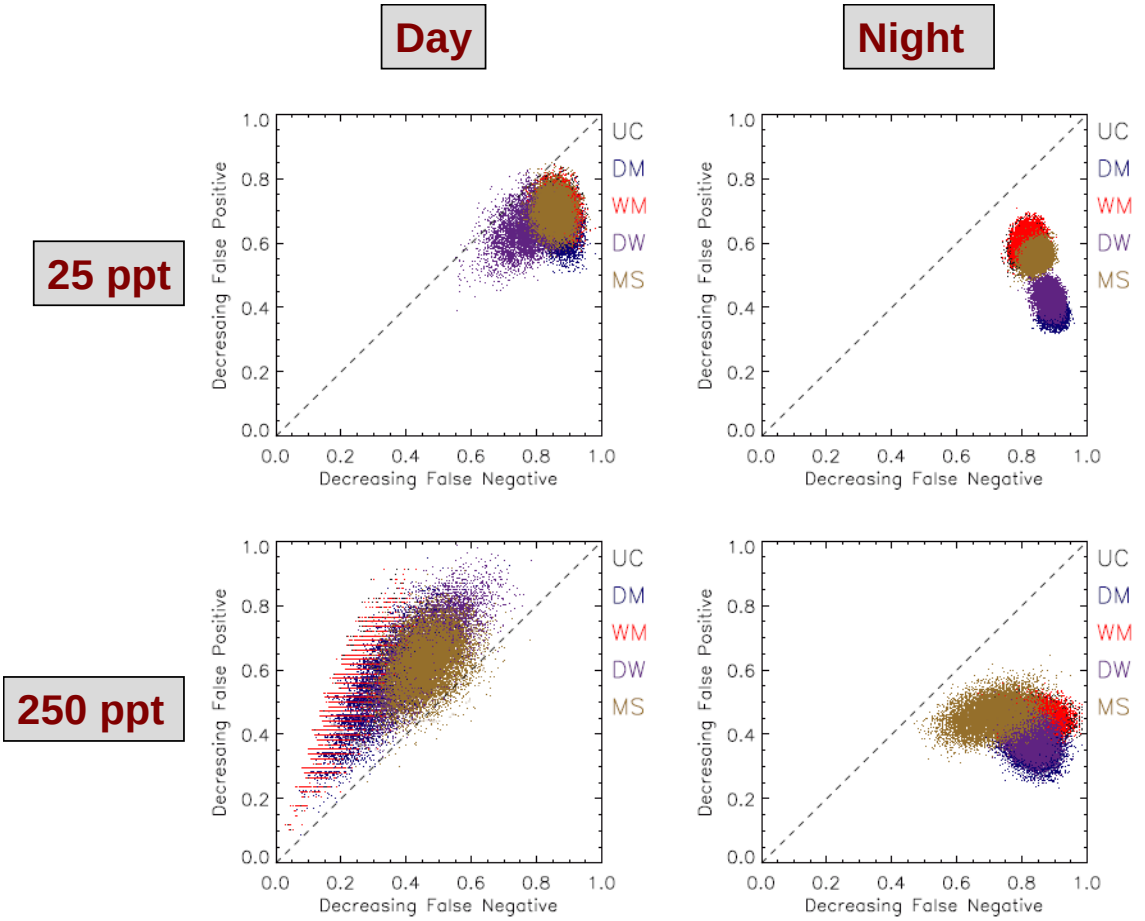


Figure 3-50. Comparisons of Threshold-Based MOE Values – 25 ppt (top) and 250 ppt (bottom) – for Urban HPAC Predictions (UC, DM, WM, DW, MS) of the 17 Daytime (left) and 12 Nighttime (right) Releases of JU03. These MOE values are for predictions of 30-minute average concentrations on the arcs using the PO7 MET input option.

The colored clusters correspond to the approximate 0.99 confidence intervals for each of the MOE estimates. The MOE point estimates lie at the approximate center of the associated cluster.

6. Summary for PO7 MET Input Option Results

Based on the above discussions and consideration of the comparisons of the five metrics, particularly the measures of scatter (NAD, NMSE, and concentration-based MOE), the qualitative results shown in Table 3-35 appear reasonable. The improvements in nighttime predictions associated with the inclusion of **MS**, **DM**, and **DW** are similar to what was found for the BAS and BRB MET input option. However, for the PO7 option, it was also found that **MS** represents a significant improvement (in terms of measures of scatter) relative to **DM** and **DW**. These PO7-based results should be thought of as consistent with an Urban HPAC usage that employs a single observation point from above an upwind building as meteorological input – a possibly operationally representative use of Urban HPAC for some specific scenarios.

Table 3-35. Urban HPAC Modes (PO7 MET) That Consistently, Across at Least Two of the Three Scatter-Related Metrics, Led to Improved Predictive Performance of JU03 Releases

Surface Sampler Region	Day	Night
CBD	mixed	(**MS**,**DM**,**DW**) /(**UC**,**WM**) and **MS** / (**DM**,**DW**)
Arcs	mixed	(**MS**,**DM**,**DW**) /(**UC**,**WM**) and **MS** / (**DM**,**DW**)

F. SUMMARY OF URBAN HPAC MODE COMPARISONS

Table 3-36 shows, for each of the five MET input options that we considered, the Urban HPAC modes that resulted in the least scatter, and as such, summarizes the comparative results discussed in the previous section. As previously discussed, when applied to observations and predictions paired in space and time, scatter-based metrics allow for the evaluation of how well the model predicted the location and timing (at least for 30-minute averages examined here) of the observations. For this reason, we consider the three scatter-evaluating metrics – NAD, NMSE, and the concentration-based MOE – as important measures of model predictive performance.

Table 3-36 identifies the Urban HPAC modes that resulted in relative (and statistically significant) improvement for the five MET input options and the four conditions (day and night, CBD and arcs) that were examined. In addition, it is noted that the BAS, BRB, and PO7 MET input options included the use of the mass-consistent terrain adjusting wind field model SWIFT (as discussed in Chapter 2), and the PNA and ACA MET input options used MC-SCIPUFF.

A few major, and relatively robust, conclusions are apparent from Table 3-36 and the previous comparative figures and tables.[9]

Table 3-36. Urban HPAC Modes, for Five MET Input Options, That Led to Improved Predictive Performance of JU03 Releases Based on Measures of Predicted/Observed Scatter (Concentration-Based MOE, NAD, and NMSE)[a]

Condition / MET Input Option	BAS (SWIFT)	BRB (SWIFT)	PO7 (SWIFT)	PNA (MC-SCIPUFF)	ACA (MC-SCIPUFF)
Day CBD	DW/DM	mixed	mixed	(UC,WM,DM,DW) /MS and DW/DM	DW/MS
Day Arcs	(MS,DW) /(UC,WM)	mixed	mixed	mixed	no differences
Night CBD	(MS,DM,DW) /(UC,WM)	(MS,DM,DW) /(UC,WM)	(MS,DM,DW) /(UC,WM) and MS / (DM,DW)	mixed	no differences
Night Arcs	(MS,DM,DW) /(UC,WM) and DM/DW	(MS,DM,DW) /(UC,WM)	(MS,DM,DW) /(UC,WM) and MS/ (DM,DW)	mixed	MS / (UC,WM,DM)

[a] The XX/YY nomenclature denotes model(s) XX had statistically significant relative improvement over model(s) YY. For this table, the designation implies that, for at least two of the three scatter-related metrics (concentration-based MOE, NAD, and NMSE), XX showed a statistically significant improvement relative to YY. The word "mixed" implies that there was not a consistent finding of one model or models over others.

1. Day vs. Night Releases and Predictions

First, a review of the previous figures, particularly the MOE-, NAD- and NMSE-related figures, reveals a substantial difference in the performance of Urban HPAC as a function of day and night. For the SWIFT-associated MET options – BAS, BRB, and PO7 – Urban HPAC predictions resulted in substantially more scatter at night than during the day, with the exception of **MS**. For the MC-SCIPUFF-associated MET input options, the scatter results were much more similar for the day and night Urban HPAC

[9] We also investigated predictions of 2-hour concentrations, and the conclusions presented here are also consistent with those examinations. This robustness of fundamental conclusions relative to the time resolution associated with the computation of metrics (e.g., 30-minute versus 2-hour) is completely consistent with past findings [Refs. 3-1 and 3-2].

predictions, with perhaps some evidence of improved performance during the day for PNA and ACA.[10] For all five MET options, daytime releases tended to be under-predicted (when considering 30-minute average concentrations at the surface samplers in the CBD and on the arcs) and nighttime releases tended to be over-predicted (e.g., review the previous FB-related figures).

2. MSS Model Performance Behavior Differs From Other Modes

With respect to the under- and over-predictions described above, the **MS** mode typically led to less under-prediction during the day *and* less over-prediction at night than the other Urban HPAC modes (with some minor exceptions where **DM** and **DW** modes were similar to **MS**). Typically, the **MS** mode resulted in the least biased predictions of the 30-minute average concentrations at the surface samplers (CBD and arcs).

3. Relative Urban HPAC Mode Performance for Nighttime Releases: MS, DM, and DW Represented Improvements

An additional important result is that for the nighttime releases, the **MS**, **DM**, and **DW** modes offer improvement over the **UC** and **WM** modes for the three MET input options that invoked SWIFT (vice MC-SCIPUFF): BAS, BEB, and PO7. This finding was true for the samplers in the CBD and for the samplers along the arcs. This result can be considered especially important because the use of SWIFT corresponds to a recommended and default mode of Urban HPAC.[11] In addition, these MET options, particularly BAS and BRB, appear to correspond to reasonably realistic and potential operational applications of Urban HPAC. We also found that adding UWM to UDM to create the **DW** mode, did not lead to substantial or consistent significant improvements relative to using UDM alone, i.e., **DM**. This result is entirely consistent with past studies of the Urban 2000 [Ref. 3-1] and MUST [Ref. 3-2] field trials. It also should be noted that the **DW** predictions (as we ran them) took approximately 80 minutes longer per release than the corresponding **DM** prediction. These results, and past findings [Refs. 3-1 and 3-2], call into question the value of including UWM, at least as we have been able to implement this feature. We also created two sets (BAS and PO7) of low

---

[10] For example, the PNA NAD values associated with the arc samplers are lower during the day than at night and the PNA- and ACA-based NMSE values are lower during the day than at night for all but the **MS** mode.

[11] We used MC-SCIPUFF (vice SWIFT) for the PNA and ACA MET options because, for some of the releases that we attempted to create, SWIFT-related software errors resulted in HPAC aborts.

resolution **DW** (and **WM**) predictions that could be run in a few minutes (per release) as described in Chapter 2.[12] We did this, essentially, for completeness – to make sure that low resolution **DW** runs would not somehow offer improved predictions. We compared predictions associated with the low and high resolution versions of **WM** and **DW** (and **MS** as described in Chapter 2). We did not find substantial, significant differences due to the lower resolution configurations (when compared with their higher resolution pairs). In fact, only very limited evidence indicated that the higher resolution predictions ever corresponded to significant improvements over the lower resolution predictions. That is, the overall fits to the observations were quite similar across the range of resolutions that were examined.

For the nighttime releases and the MC-SCIPUFF-associated MET input options – PNA and ACA – results were mixed with no Urban HPAC mode consistently offering improvement although the **MS** mode did so for the ACA MET option (at least relative to **UC**, **WM**, and **DM**). Differences in model performance are directly examined, albeit briefly, in the next section (Section G).

4. Relative Urban HPAC Mode Performance for Daytime Releases Was Mixed and Inconsistent

For the daytime releases, no consistent trend was found. For example, for the BAS-associated predictions on the arcs, the **MS** and **DW** modes offer improvement (e.g., less scatter) over the **UC** and **WM** modes but, for the PNA-associated predictions in the CBD, the **UC**, **WM**, and **DW** resulted in improved scatter relative to the **MS** mode. However, in the latter PNA-based case, the observed improvements in scatter for the **UC**, **WM**, and **DW** predictions come at the cost of a large under-prediction relative to **MS** (Figure 3-21). That is, even the statistically detectable difference in scatter metrics for this case may be of questionable value.

5. Concentration-Based Versus Threshold-Based MOE Values

Predictions of exceeding a relatively low concentration threshold (5 × and 50 × background) were more accurate (as measured by the MOE) than predictions of absolute 30-minute average concentrations. For example, for the BAS MET option and the **MS** mode, the day and night concentration-based MOE values for the CBD are about (0.59, 0.42) and (0.53, 0.34), respectively (Figure 3-1). The corresponding day/night 25 and

---

[12] Additional discussion of these "run time" analysis effort can be found in Ref. 3-4.

250 ppt MOE values are (0.90, 0.72) / (0.77, 0.67) and (0.92, 0.67) / (0.72, 0.61), respectively (Figure 3-9). These MOE results, and similar ones found for the other Urban HPAC mode/MET option combinations, indicate substantial improvements when the thresholds are examined. This result is entirely consistent with past studies of the Urban 2000 [Ref. 3-1] and MUST [Ref. 3-2] field trials. An important implication of the above finding is that using Urban HPAC to predict the extent (in time and space) to which a relatively low threshold is (or might be) exceeded is likely to lead to a more accurate representation of a hazardous release (or area) than using Urban HPAC to predict the actual concentrations in time and space (e.g., perhaps needed for a detailed and complete assessment of expected casualties given a human effects model that requires concentration-time histories as a function of location).

Appendix B provides supplementary plots that compare MOE values for the 25 sets of Urban HPAC predictions during the day and at night, in the CBD and on the arcs, and for concentration-based, 25 ppt threshold-based, and 250 ppt threshold-based calculations.

G. BRIEF COMPARISON ON MET INPUT OPTIONS

Figures 3-51 and 3-52 provide comparisons of FB values for Urban HPAC predictions of the daytime (red) and nighttime (blue) releases of JU03 within the CBD (Figure 3-51) and along the arcs (Figure 3-52) using the five MET input options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB). As has been previously discussed, day releases were generally under-predicted and the night releases were generally over-predicted both within the CBD and on the arcs. The **MS** mode represented an exception to the above, particularly for the CBD samplers (Figure 3-51).

Figures 3-51 and 3-52 also indicate a difference between the two MC-SCIPUFF-associated MET options – PNA and ACA – and the SWIFT-associated MET options – BAS, BRB, and PO7. For the **UC**, **WM**, **DM**, and **DW** modes, the PNA and ACA options resulted in less material being predicted at the surface samplers, for example, less over-prediction at night and more under-prediction during the day relative to the other MET options. This type of behavior could be caused by a faster wind speed being associated with the MC-SCIPUFF options relative to the SWIFT options. In this case, the faster wind speed simply blows material past the samplers (e.g., in the CBD) too quickly and results in less material predicted in the 30-minute averages.

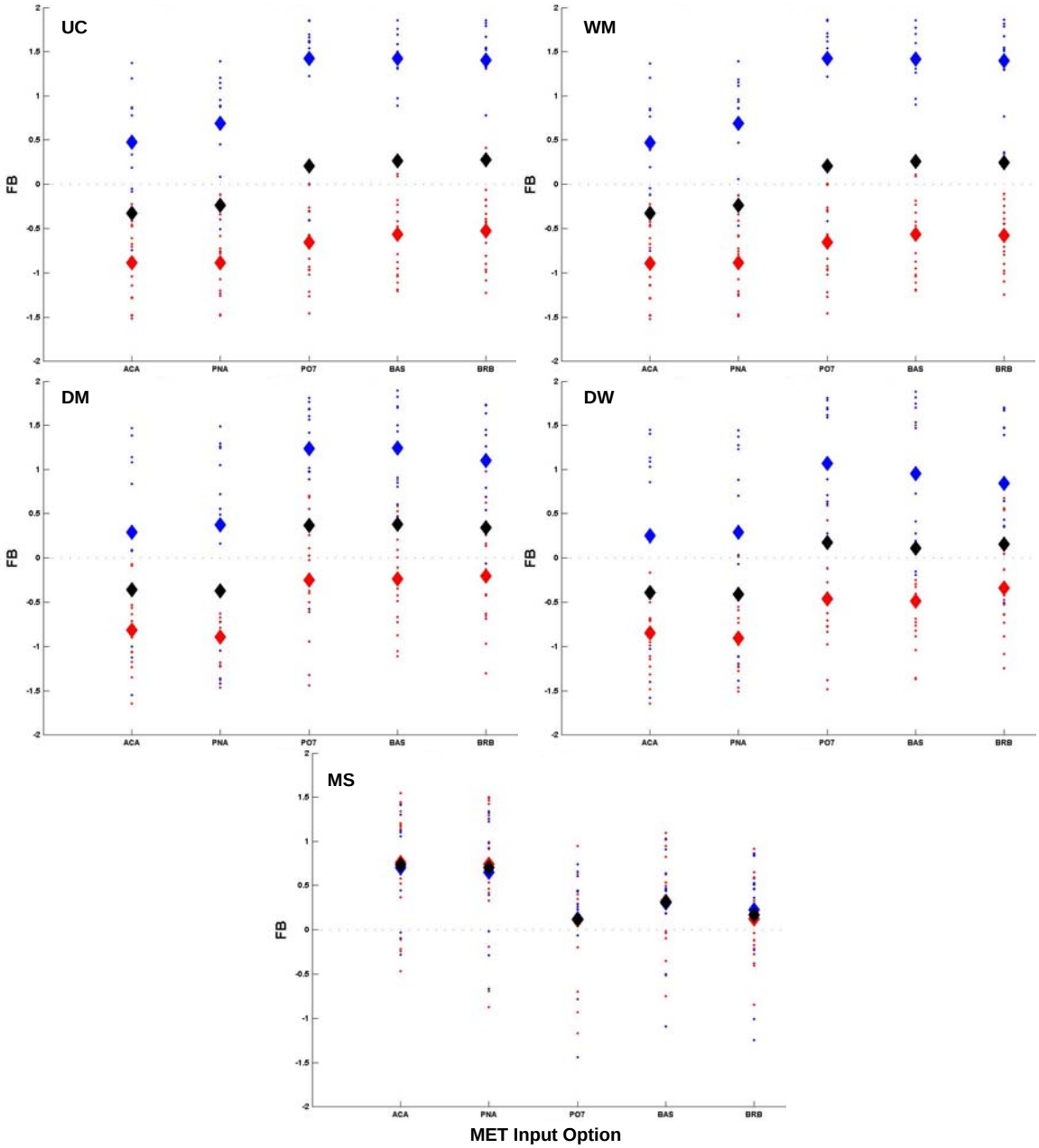


Figure 3-51. Comparisons of FB Values for Urban HPAC Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 Within the CBD Using the Five MET Input Options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB)

The smaller colored points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB values (day = red and night = blue) with the large black diamond representing the overall average for all 29 releases.

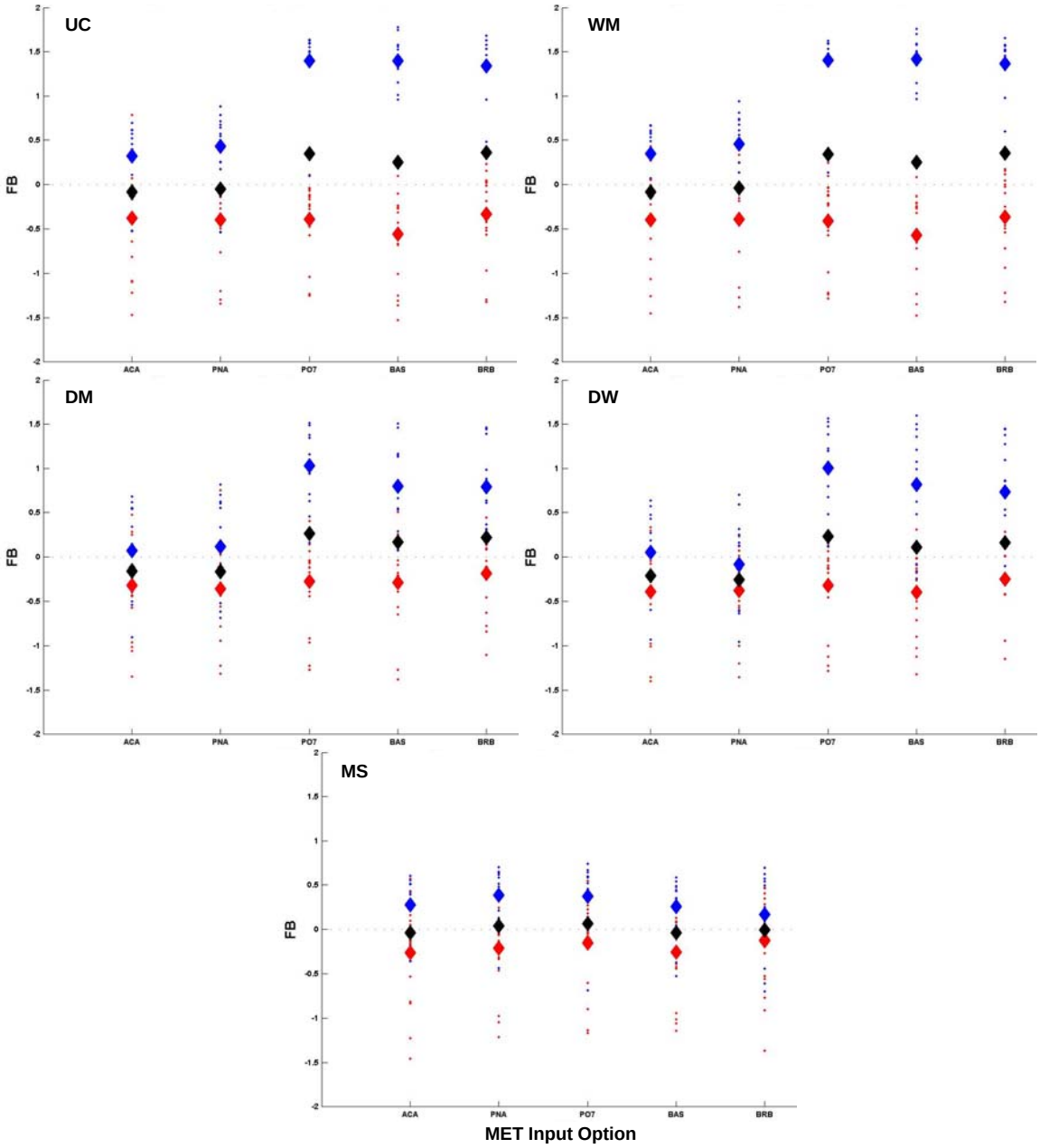


Figure 3-52. Comparisons of FB Values for Urban HPAC Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 Along the Arcs Using the Five MET Input Options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB)

The smaller colored points correspond to FB values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB values (day = red and night = blue) with the large black diamond representing the overall average for all 29 releases.

This faster wind speed hypothesis was supported by examinations of contour plots. Figure 3-53 shows an example that compares predictions, using the BRB and ACA MET input options for two releases – one day (IOP 5, release1) and one night (IOP 7, Release 3). The below contours are consistent with the notion that the "plume" moved through the sampling region much quicker for the ACA option relative to the BAS option. For example, note the relative amounts of material remaining on the samplers during the second 30-minute period for IOP 5, Release 1 and the third 30-minute period for IOP 7, Release 3 for the two modes – BRB and ACA.

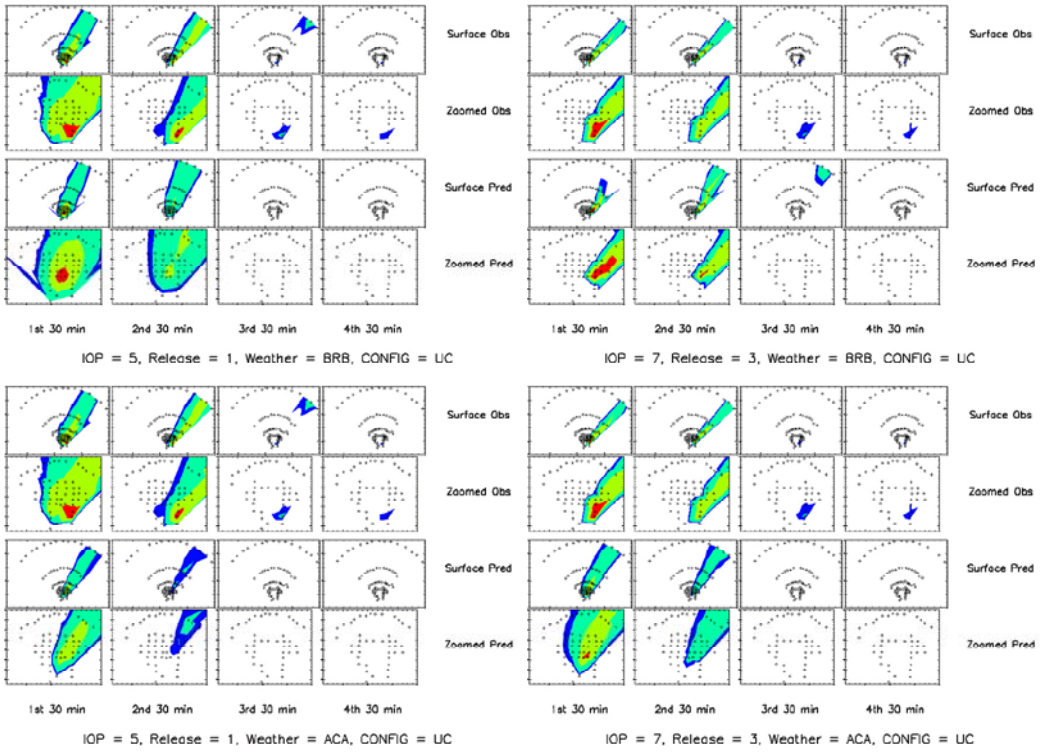


Figure 3-53. Comparison of BRB and ACA MET Option Predictions of 30-Minute Average Concentrations

Contours for the four 30-minute time periods that were monitored during IOP 5, Release 1 (left) and IOP 7, Release 3 (right) for the surface observations and predictions at the surface samplers are shown for all surface samplers and for the CBD-"Zoomed." These predictions are based on the UC mode of Urban HPAC.

In Figure 3-53 the ACA-based predicted plume also appears to move through the sampling area quicker than the "observed" plume. One can imagine how such an effect, on average over the 29 releases, would lead to less material predicted at the samplers for ACA-based predictions relative to BRB-based predictions. This type of finding occurred often when comparing the two MC-SCIPUFF-based MET options to the three SWIFT-based options. This effect could account for the finding described above for Figures 3-51

and 3-52 of "less over-prediction at night and more under-prediction during the day relative to the other MET options" for the MC-SCIPUFF-based options (PNA and ACA). In the case of the nighttime predictions, in which concentrations were typically over-predicted, such an error in wind speed (too fast) would partially compensate for the underlying over-prediction and lead to "better" predictions (at least in terms of FB).

Another interpretation of this observation is that some other combination of problems, perhaps including SWIFT-based winds that are too slow, could be the ultimate cause. Considerations of fundamental differences between SWIFT and MC-SCIPUFF, and how such differences may account for the above findings, are an ongoing research topic for our group.

Comparisons of NAD values for Urban HPAC predictions of the daytime (red) and nighttime (blue) releases of JU03 are shown in Figures 3-54 (CBD) and 3-55 (arcs). It can be seen that for most Urban HPAC modes, the predictions of night releases result in more scatter than those associated with the day releases (as has been previously discussed). NMSE, BNMSE, and NAD values (all measures of scatter) for 25 sets of Urban HPAC predictions of JU03 for day and night, and for CBD and along the arcs are listed in Table 3-37.[13] The underlined values in Table 3-37 represent the best (least scatter) value for each MET input option and the boldfaced underlined values correspond to the best values for all 25 sets of predictions for each scatter metric. Tables 3-38 and 3-39 list the NAD and NMSE values, respectively, from least scatter (best) to most scatter (worst). **MS**-based predictions resulted in the least scatter in seven of the eight categories (day/night, CBD/arcs, and NAD/NMSE). The PO7_**DW** is the sole exception to the above, having the best NAD value for the day-CBD condition. The best NAD and NMSE values by conditions are shown below.

- Day, CBD: PO7_**DW** (NAD = 0.42) and PO7_**MS** (NMSE = 16.91)
- Day Arcs: PO7_ **MS** (NAD = 0.49) and PO7_**MS** (NMSE = 27.68)
- Night CBD: PNA_ **MS** (NAD = 0.31) and PNA_**MS** (NMSE = 7.69)
- Night Arcs: ACA_ **MS** (NAD = 0.35) and ACA_**MS** (NMSE = 9.49)

Tables 3-40 and 3-41 list NAD and NMSE values, respectively, by Urban HPAC mode and with each MET option (for each mode) ordered from least to most scatter. The

---

[13] In addition to the NAD and NMSE measures of scatter, BNMSE values (described in Chapter 2) are presented in Table 3-37. Comparative results associated with BNMSE and NAD are very similar, and hence BNMSE is not further discussed in the tables that follow.

underlined MET options correspond to those predictions that were found to be statistically significantly different (p-value < 0.05) from the first option listed. Several conclusions associated with comparisons of MET options and based on scatter (and concentration-based MOE values not shown here) are summarized below.

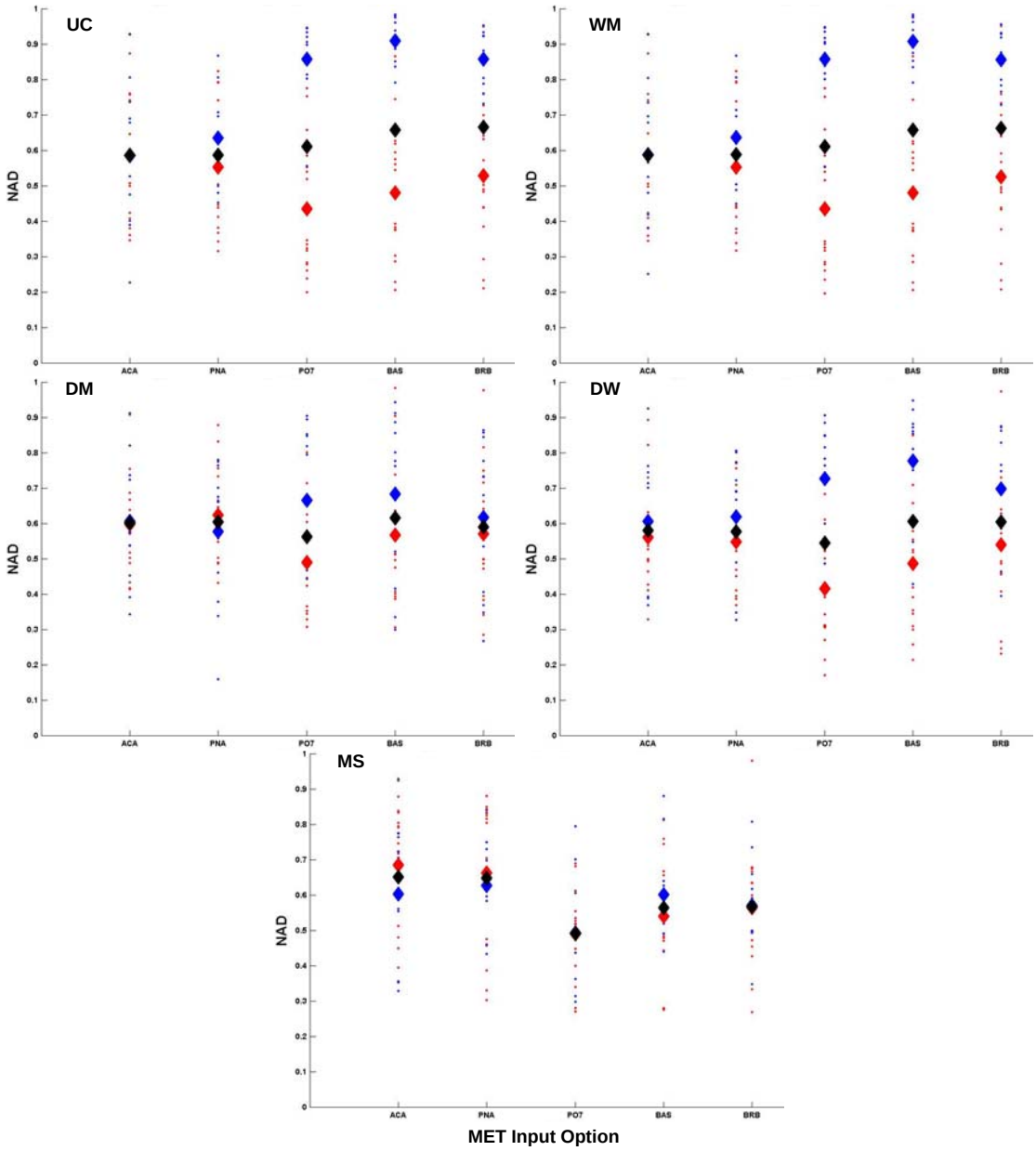


Figure 3-54. Comparisons of NAD Values for Urban HPAC Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 Within the CBD Using the Five MET Input Options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB)

The smaller colored points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD values (day = red and night = blue) with the large black diamond representing the overall average for all 29 releases.

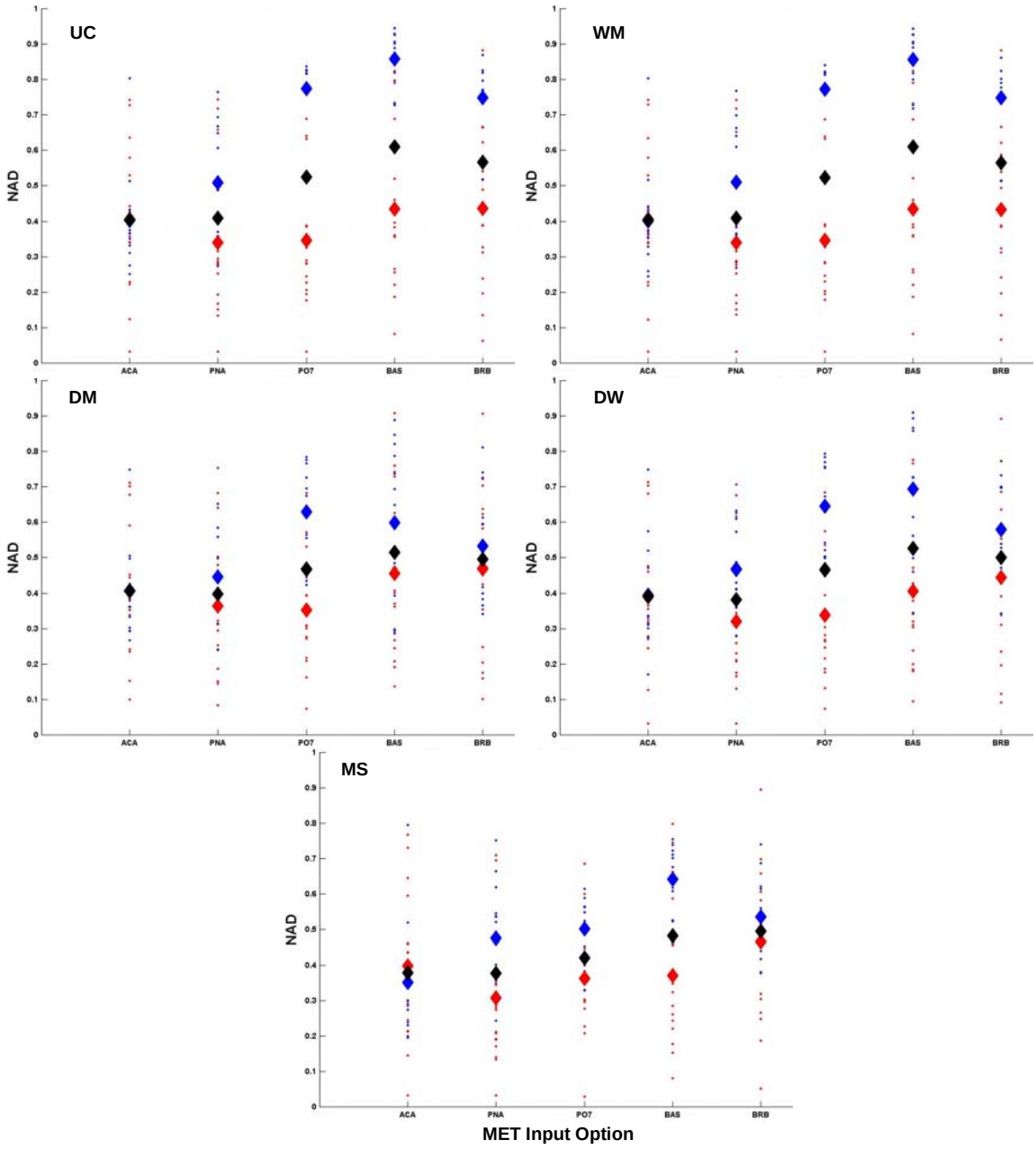


Figure 3-55. Comparisons of NAD Values for Urban HPAC Predictions of the Daytime (red) and Nighttime (blue) Releases of JU03 Along the Arcs Using the Five MET Input Options (labeled along the x-axes of each chart as ACA, PNA, PO7, BAS, and BRB)

The smaller colored points correspond to NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average NAD values (day = red and night = blue) with the large black diamond representing the overall average for all 29 releases.

Table 3-37. NMSE, BNMSE, and NAD Values for 25 Sets of Urban HPAC Predictions of JU03. Values are shown for day and night and for CBD and Arcs. The underlined values represent the best (least scatter) value for each MET input option (e.g., DW for the PNA option and NMSE during the day and in the CBD) and the boldfaced underlined values correspond to the best values for all 25 sets of predictions for each scatter metric.

		Day	CBD		Night	CBD		Day	Arcs		Night	Arcs	
Mode	MET	NMSE	BNMSE	NAD	NMSE	BNMSE	NAD	NMSE	BNMSE	NAD	NMSE	BNMSE	NAD
UC	PO7	29.33	0.30	0.44	506.26	0.81	0.86	11.76	0.23	0.35	143.53	0.71	0.77
WM	PO7	29.27	0.30	0.44	521.59	0.81	0.86	11.80	0.23	0.35	141.98	0.71	0.77
DM	PO7	22.70	0.34	0.49	143.00	0.58	0.67	10.47	0.22	0.35	35.54	0.44	0.63
DW	PO7	20.08	**0.27**	**0.42**	159.87	0.65	0.73	11.27	0.21	0.34	49.52	0.51	0.64
MS	PO7	**16.91**	0.30	0.49	**27.68**	**0.27**	**0.49**	9.34	0.22	0.36	14.86	0.34	0.50
UC	PNA	45.22	0.43	0.55	60.41	0.56	0.64	11.54	0.23	0.34	14.43	0.34	0.51
WM	PNA	45.09	0.43	0.55	59.63	0.56	0.64	11.41	0.22	0.34	14.65	0.34	0.51
DM	PNA	38.81	0.53	0.62	78.59	0.51	0.58	11.80	0.25	0.36	12.47	0.25	0.45
DW	PNA	33.17	0.43	0.55	91.96	0.57	0.62	9.40	0.19	0.32	14.09	0.30	0.47
MS	PNA	119.28	0.60	0.66	63.36	0.54	0.63	**7.69**	**0.18**	**0.31**	11.58	0.29	0.48
UC	ACA	50.18	0.46	0.59	71.10	0.52	0.59	17.04	0.30	0.41	14.71	0.25	0.40
WM	ACA	50.36	0.47	0.59	71.02	0.52	0.59	17.13	0.30	0.41	14.47	0.25	0.40
DM	ACA	38.54	0.49	0.60	96.66	0.53	0.61	12.88	0.29	0.41	13.26	0.24	0.41
DW	ACA	33.97	0.44	0.56	90.38	0.51	0.61	15.30	0.28	0.39	12.27	0.23	0.39
MS	ACA	100.40	0.61	0.69	77.88	0.53	0.60	15.57	0.29	0.40	**9.49**	**0.19**	**0.35**
UC	BAS	27.22	0.33	0.48	346.48	0.90	0.91	19.18	0.37	0.44	162.43	0.86	0.86
WM	BAS	27.15	0.33	0.48	340.29	0.90	0.91	19.17	0.37	0.44	161.63	0.86	0.86
DM	BAS	31.48	0.44	0.57	185.90	0.63	0.68	13.42	0.37	0.46	24.96	0.47	0.60
DW	BAS	17.68	0.34	0.49	218.92	0.74	0.78	13.07	0.32	0.41	50.34	0.63	0.69
MS	BAS	24.37	0.39	0.54	28.58	0.42	0.60	10.29	0.26	0.37	18.32	0.53	0.64
UC	BRB	29.07	0.42	0.53	488.84	0.81	0.86	13.45	0.36	0.44	120.48	0.68	0.75
WM	BRB	29.66	0.42	0.53	493.54	0.81	0.86	13.62	0.36	0.43	121.84	0.68	0.75
DM	BRB	21.08	0.43	0.57	133.79	0.54	0.62	11.85	0.40	0.47	26.17	0.35	0.53
DW	BRB	20.61	0.40	0.54	150.00	0.63	0.70	12.11	0.37	0.45	34.99	0.44	0.58
MS	BRB	17.94	0.38	0.57	33.38	0.37	0.57	17.94	0.38	0.57	33.38	0.37	0.57

Table 3-38. NAD Values for 25 Sets of Urban HPAC Predictions of JU03 Ordered from Least to Most Scatter. Values are shown for day and night and for CBD and Arcs.

	Day	CBD	Metric	Night	CBD	Metric	Day	Arcs	Metric	Night	Arcs	Metric
Rank	Mode	MET	NAD	Mode	MET	NAD	Mode	MET	NAD	Mode	MET	NAD
1	DW	PO7	0.42	MS	PO7	0.49	MS	PNA	0.31	MS	ACA	0.35
2	WM	PO7	0.44	MS	BRB	0.57	DW	PNA	0.32	DW	ACA	0.39
3	UC	PO7	0.44	DM	PNA	0.58	DW	PO7	0.34	WM	ACA	0.40
4	WM	BAS	0.48	UC	ACA	0.59	WM	PNA	0.34	UC	ACA	0.40
5	UC	BAS	0.48	WM	ACA	0.59	UC	PNA	0.34	DM	ACA	0.41
6	DW	BAS	0.49	MS	BAS	0.60	UC	PO7	0.35	DM	PNA	0.45
7	DM	PO7	0.49	MS	ACA	0.60	WM	PO7	0.35	DW	PNA	0.47
8	MS	PO7	0.49	DM	ACA	0.61	DM	PO7	0.35	MS	PNA	0.48
9	WM	BRB	0.53	DW	ACA	0.61	MS	PO7	0.36	MS	PO7	0.50
10	UC	BRB	0.53	DM	BRB	0.62	DM	PNA	0.36	UC	PNA	0.51
11	MS	BAS	0.54	DW	PNA	0.62	MS	BAS	0.37	WM	PNA	0.51
12	DW	BRB	0.54	MS	PNA	0.63	DW	ACA	0.39	DM	BRB	0.53
13	DW	PNA	0.55	UC	PNA	0.64	MS	ACA	0.40	MS	BRB	0.57
14	WM	PNA	0.55	WM	PNA	0.64	DM	ACA	0.41	DW	BRB	0.58
15	UC	PNA	0.55	DM	PO7	0.67	UC	ACA	0.41	DM	BAS	0.60
16	DW	ACA	0.56	DM	BAS	0.68	DW	BAS	0.41	DM	PO7	0.63
17	MS	BRB	0.57	DW	BRB	0.70	WM	ACA	0.41	MS	BAS	0.64
18	DM	BAS	0.57	DW	PO7	0.73	WM	BRB	0.43	DW	PO7	0.64
19	DM	BRB	0.57	DW	BAS	0.78	WM	BAS	0.44	DW	BAS	0.69
20	UC	ACA	0.59	WM	BRB	0.86	UC	BAS	0.44	UC	BRB	0.75
21	WM	ACA	0.59	UC	BRB	0.86	UC	BRB	0.44	WM	BRB	0.75
22	DM	ACA	0.60	WM	PO7	0.86	DW	BRB	0.45	WM	PO7	0.77
23	DM	PNA	0.62	UC	PO7	0.86	DM	BAS	0.46	UC	PO7	0.77
24	MS	PNA	0.66	WM	BAS	0.91	DM	BRB	0.47	WM	BAS	0.86
25	MS	ACA	0.69	UC	BAS	0.91	MS	BRB	0.57	UC	BAS	0.86

Table 3-39. NMSE Values for 25 Sets of Urban HPAC Predictions of JU03 Ordered from Least to Most Scatter. Values are shown for day and night and for CBD and Arcs.

	Day	CBD	Metric	Night	CBD	Metric	Day	Arcs	Metric	Night	Arcs	Metric
Rank	Mode	MET	NMSE	Mode	MET	NMSE	Mode	MET	NMSE	Mode	MET	NMSE
1	MS	PO7	16.91	MS	PO7	27.68	MS	PNA	7.69	MS	ACA	9.49
2	DW	BAS	17.68	MS	BAS	28.58	MS	PO7	9.34	MS	PNA	11.58
3	MS	BRB	17.94	MS	BRB	33.38	DW	PNA	9.40	DW	ACA	12.27
4	DW	PO7	20.08	WM	PNA	59.63	MS	BAS	10.29	DM	PNA	12.47
5	DW	BRB	20.61	UC	PNA	60.41	DM	PO7	10.47	DM	ACA	13.26
6	DM	BRB	21.08	MS	PNA	63.36	DW	PO7	11.27	DW	PNA	14.09
7	DM	PO7	22.70	WM	ACA	71.02	WM	PNA	11.41	UC	PNA	14.43
8	MS	BAS	24.37	UC	ACA	71.10	UC	PNA	11.54	WM	ACA	14.47
9	WM	BAS	27.15	MS	ACA	77.88	UC	PO7	11.76	WM	PNA	14.65
10	UC	BAS	27.22	DM	PNA	78.59	DM	PNA	11.80	UC	ACA	14.71
11	UC	BRB	29.07	DW	ACA	90.38	WM	PO7	11.80	MS	PO7	14.86
12	WM	PO7	29.27	DW	PNA	91.96	DM	BRB	11.85	MS	BAS	18.32
13	UC	PO7	29.33	DM	ACA	96.66	DW	BRB	12.11	DM	BAS	24.96
14	WM	BRB	29.66	DM	BRB	133.79	DM	ACA	12.88	DM	BRB	26.17
15	DM	BAS	31.48	DM	PO7	143.00	DW	BAS	13.07	MS	BRB	33.38
16	DW	PNA	33.17	DW	BRB	150.00	DM	BAS	13.42	DW	BRB	34.99
17	DW	ACA	33.97	DW	PO7	159.87	UC	BRB	13.45	DM	PO7	35.54
18	DM	ACA	38.54	DM	BAS	185.90	WM	BRB	13.62	DW	PO7	49.52
19	DM	PNA	38.81	DW	BAS	218.92	DW	ACA	15.30	DW	BAS	50.34
20	WM	PNA	45.09	WM	BAS	340.29	MS	ACA	15.57	UC	BRB	120.48
21	UC	PNA	45.22	UC	BAS	346.48	UC	ACA	17.04	WM	BRB	121.84
22	UC	ACA	50.18	UC	BRB	488.84	WM	ACA	17.13	WM	PO7	141.98
23	WM	ACA	50.36	WM	BRB	493.54	MS	BRB	17.94	UC	PO7	143.53
24	MS	ACA	100.40	UC	PO7	506.26	WM	BAS	19.17	WM	BAS	161.63
25	MS	PNA	119.28	WM	PO7	521.59	UC	BAS	19.18	UC	BAS	162.43

Table 3-40. NAD Values for 25 Sets of Urban HPAC Predictions of JU03 Separated by Mode and Ordered from Least to Most Scatter. Values are shown for day and night and for CBD and Arcs. The underlined MET options correspond to those predictions that were found to be statistically significantly different (p-value < 0.05) from the first option listed (e.g., for the day-CBD, the PO7_DM predictions represented a detectable improvement relative to BAS_DM, BRB_DM, ACA_DM, and PNA_DM.)

Day	CBD	Metric	Night	CBD	Metric	Day	Arcs	Metric	Night	Arcs	Metric
Mode	MET	NAD	Mode	MET	NAD	Mode	MET	NAD	Mode	MET	NAD
DM	PO7	0.49	DM	PNA	0.58	DM	PO7	0.35	DM	ACA	0.41
DM	BAS	0.57	DM	ACA	0.61	DM	PNA	0.36	DM	PNA	0.45
DM	BRB	0.57	DM	BRB	0.62	DM	ACA	0.41	DM	BRB	0.53
DM	ACA	0.60	DM	PO7	0.67	DM	BAS	0.46	DM	BAS	0.60
DM	PNA	0.62	DM	BAS	0.68	DM	BRB	0.47	DM	PO7	0.63
DW	PO7	0.42	DW	ACA	0.61	DW	PNA	0.32	DW	ACA	0.39
DW	BAS	0.49	DW	PNA	0.62	DW	PO7	0.34	DW	PNA	0.47
DW	BRB	0.54	DW	BRB	0.70	DW	ACA	0.39	DW	BRB	0.58
DW	PNA	0.55	DW	PO7	0.73	DW	BAS	0.41	DW	PO7	0.64
DW	ACA	0.56	DW	BAS	0.78	DW	BRB	0.45	DW	BAS	0.69
MS	PO7	0.49	MS	PO7	0.49	MS	PNA	0.31	MS	ACA	0.35
MS	BAS	0.54	MS	BRB	0.57	MS	PO7	0.36	MS	PNA	0.48
MS	BRB	0.57	MS	BAS	0.60	MS	BAS	0.37	MS	PO7	0.50
MS	PNA	0.66	MS	ACA	0.60	MS	ACA	0.40	MS	BRB	0.57
MS	ACA	0.69	MS	PNA	0.63	MS	BRB	0.57	MS	BAS	0.64
UC	PO7	0.44	UC	ACA	0.59	UC	PNA	0.34	UC	ACA	0.40
UC	BAS	0.48	UC	PNA	0.64	UC	PO7	0.35	UC	PNA	0.51
UC	BRB	0.53	UC	BRB	0.86	UC	ACA	0.41	UC	BRB	0.75
UC	PNA	0.55	UC	PO7	0.86	UC	BAS	0.44	UC	PO7	0.77
UC	ACA	0.59	UC	BAS	0.91	UC	BRB	0.44	UC	BAS	0.86
WM	PO7	0.44	WM	ACA	0.59	WM	PNA	0.34	WM	ACA	0.40
WM	BAS	0.48	WM	PNA	0.64	WM	PO7	0.35	WM	PNA	0.51
WM	BRB	0.53	WM	BRB	0.86	WM	ACA	0.41	WM	BRB	0.75
WM	PNA	0.55	WM	PO7	0.86	WM	BRB	0.43	WM	PO7	0.77
WM	ACA	0.59	WM	BAS	0.91	WM	BAS	0.44	WM	BAS	0.86

Table 3-41. NMSE Values for 25 Sets of Urban HPAC Predictions of JU03 Separated by Mode and Ordered from Least to Most Scatter. Values are shown for day and night and for CBD and Arcs. The underlined MET options correspond to those predictions that were found to be statistically significantly different (p-value < 0.05) from the first option listed (e.g., for the day-CBD, the BRB_DM predictions represented a detectable improvement relative to ACA_DM and PNA_DM.)

Day	CBD	Metric	Night	CBD	Metric	Day	Arcs	Metric	Night	Arcs	Metric
Mode	MET	NMSE	Mode	MET	NMSE	Mode	MET	NMSE	Mode	MET	NMSE
DM	BRB	21.08	DM	PNA	78.59	DM	PO7	10.47	DM	PNA	12.47
DM	PO7	22.70	DM	ACA	96.66	DM	PNA	11.80	DM	ACA	13.26
DM	BAS	31.48	DM	BRB	133.79	DM	BRB	11.85	DM	BAS	24.96
DM	ACA	38.54	DM	PO7	143.00	DM	ACA	12.88	DM	BRB	26.17
DM	PNA	38.81	DM	BAS	185.90	DM	BAS	13.42	DM	PO7	35.54
DW	BAS	17.68	DW	ACA	90.38	DW	PNA	9.40	DW	ACA	12.27
DW	PO7	20.08	DW	PNA	91.96	DW	PO7	11.27	DW	PNA	14.09
DW	BRB	20.61	DW	BRB	150.00	DW	BRB	12.11	DW	BRB	34.99
DW	PNA	33.17	DW	PO7	159.87	DW	BAS	13.07	DW	PO7	49.52
DW	ACA	33.97	DW	BAS	218.92	DW	ACA	15.30	DW	BAS	50.34
MS	PO7	16.91	MS	PO7	27.68	MS	PNA	7.69	MS	ACA	9.49
MS	BRB	17.94	MS	BAS	28.58	MS	PO7	9.34	MS	PNA	11.58
MS	BAS	24.37	MS	BRB	33.38	MS	BAS	10.29	MS	PO7	14.86
MS	ACA	100.40	MS	PNA	63.36	MS	ACA	15.57	MS	BAS	18.32
MS	PNA	119.28	MS	ACA	77.88	MS	BRB	17.94	MS	BRB	33.38
UC	BAS	27.22	UC	PNA	60.41	UC	PNA	11.54	UC	PNA	14.43
UC	BRB	29.07	UC	ACA	71.10	UC	PO7	11.76	UC	ACA	14.71
UC	PO7	29.33	UC	BAS	346.48	UC	BRB	13.45	UC	BRB	120.48
UC	PNA	45.22	UC	BRB	488.84	UC	ACA	17.04	UC	PO7	143.53
UC	ACA	50.18	UC	PO7	506.26	UC	BAS	19.18	UC	BAS	162.43
WM	BAS	27.15	WM	PNA	59.63	WM	PNA	11.41	WM	ACA	14.47
WM	PO7	29.27	WM	ACA	71.02	WM	PO7	11.80	WM	PNA	14.65
WM	BRB	29.66	WM	BAS	340.29	WM	BRB	13.62	WM	BRB	121.84
WM	PNA	45.09	WM	BRB	493.54	WM	ACA	17.13	WM	PO7	141.98
WM	ACA	50.36	WM	PO7	521.59	WM	BAS	19.17	WM	BAS	161.63

- During the day in the CBD, the PO7 MET input option resulted in the best NAD values for all Urban HPAC modes. For the NMSE comparisons, the results are somewhat mixed with BRB, BAS, and PO7 leading to the best values depending on the Urban HPAC mode. However, the overall best NMSE value is associated with PO7_**MS** ($NMSE_{PO7\_MS} = 16.91$).
- At night in the CBD, the PNA and ACA MET input options resulted in the best NAD and NMSE values for four of the Urban HPAC modes. This may partially be due to the compensating errors associated with slower speeds through the CBD for these MC-SCIPUFF driven options (as previously discussed). The exception to the above is the **MS** mode, which led to the best overall NAD and NMSE values for nighttime predictions in the CBD – $NAD_{PO7\_MS} = 0.49$ and $NMSE_{PO7\_MS} = 27.68$.
- During the day on the arcs, the PNA MET input options resulted in the best NAD and NMSE values for four of the Urban HPAC modes. The exception was the PO7_**DM** mode for both NAD and NMSE.

- At night on the arcs, the PNA and ACA MET input options resulted in the best NAD and NMSE values for all five Urban HPAC modes, and the ACA mode was best for all five modes when only NAD was considered. This may partially be due to the compensating errors discussed earlier (but this is still under investigation).
- Based on Figures 3-40 through 3-41 (and qualitatively), the following MET input options performed best (in terms of least scatter):
  - Day, CBD: NAD ⇒ PO7; NMSE ⇒ mixed (3 BAS, 1 BRB, 1 PO7)
  - Night, CBD: PNA and ACA; PO7 exception with **MS**
  - Day, Arcs: PNA; PO7 exception with **DM**
  - Night, Arcs: NAD ⇒ ACA and NMSE ⇒ PNA and ACA

H. COMPARISON OF JU03 (OKLAHOMA CITY) AND PREVIOUS FINDINGS: URBAN 2000 (SALT LAKE CITY) AND MUST

The releases associated with the Urban 2000 and MUST field experiments were conducted at night or in the very early morning while atmospheric conditions were still relatively stable. Since there were substantial differences in Urban HPAC model behavior and performance as a function of day and night, the most appropriate JU03 results for comparison to Urban 2000 and MUST are those associated with the nighttime releases. Similarly, previous studies included the **UC**, **WM**, **DM**, and **DW** modes but not the **MS** mode, because it was not available at those times. Therefore, for these comparisons to previous studies we do not consider the more recent **MS** JU03 predictions. The previously created Urban 2000 and MUST predictions included the use of SWIFT in all cases. In fact, releases that led to SWIFT-generated errors were not considered in these previous studies – that is, no MC-SCIPUFF-based predictions were examined in the previous studies. Given the above considerations, the previously discussed JU03 findings that are particularly consistent with previous Urban 2000 and MUST conclusions include:

- In general, Urban HPAC modes led to over-predictions of the surface sampler concentrations. For JU03, the median nighttime FB value (not considering the **MS** mode) for the CBD and the arcs was 1.01 and 0.79 (or over-prediction factors of about 3.0 and 2.3), respectively. For Urban 2000, the comparable FB values (for the CBD and on the arcs, respectively) were 0.48 (over-prediction factor of about 1.6) and 0.70 (over-prediction factor of about 2.1). For MUST, over-

prediction factors, across the entire 200-meter square of samplers, varied by Urban HPAC mode from about 1.0 to 3.0.

- We also found for all three field experiments (JU03, Urban 2000, and MUST) that adding UWM to UDM to create the **DW** mode, did not lead to *substantial or consistent* significant improvements relative to using UDM alone, i.e., **DM**.
- For all three field trial studies, predictions of exceeding a relatively low concentration threshold were more accurate (as measured by the MOE) than predictions of absolute 30-minute average concentrations.

With respect to NAD, the JU03 Urban HPAC predictions generally resulted in less scatter than the comparable Urban 2000 predictions. For example for Urban 2000, the downtown and arc-based NAD values were never less than 0.60 for any of the 20 sets of predictions that were examined and when considering 30-minute average concentrations. As can be seen in Table 3-38, a few sets of Urban HPAC JU03 nighttime predictions (after excluding **MS** for these comparisons) resulted in NAD values less than 0.60, including SWIFT-based values on the arcs at night of 0.53 and 0.58 for the BRB-based **DM** and **DW** predictions, respectively. For MUST, NAD values are not easily comparable because the time resolutions that were examined – 10, 60, and 300 seconds; and about 15 minutes [Ref. 3-6] – are substantially different from those examined during JU03 and Urban 2000. Nonetheless, for the ≈15 minute time resolution, NAD values for the 20 sets of Urban HPAC MUST predictions were never less than 0.45, and typically above about 0.55. The relative improvement in model fit (less scatter as measured by NAD) for the JU03 predictions relative to those of Urban 2000 could be partially explained by improved Urban HPAC models available since the time of the Urban 2000 study *and/or* improved MET inputs used in this JU03 study – i.e., MET inputs that better represent the actual winds that transport the plume.

The FACx metric (x = 2, 5, and 10, as in Chapter 2) considers the ratios between the predictions and the observations at each point in space and time – here for 30-minute average concentrations. This metric is equally sensitive to the smaller and larger concentrations, whereas metrics such as NAD, NMSE, and the concentration-based MOE can be dominated by the larger concentrations. As such, FACx can be particularly sensitive to mismatches in plume transport direction (e.g., several samplers with relatively small observed concentrations missed on one side of the plume because of a 10-20 degree transport error could easily lead to predictions and observations differing by more than a factor of two).

The FACx metric can also be very sensitive to the underlying sampling distribution and the data processing protocols. For comparisons between different field experiments, we followed consistent protocols in order to allow for fair comparisons between field trials. We examined several data protocol options for our examinations of FACx. First, we considered all sampler locations where either the prediction or observation was at least background + 1 ppt. This was done to ensure that FACx values were not overly influenced by samplers where both the observation and prediction were simply defined as background – 5 ppt. The assumed backgrounds for JU03 and Urban 2000 [Ref. 3-1] were 5 and 3 ppt, respectively. Next we computed results by excluding all comparisons in which both the prediction *and* observation were below 15 ppt.[14] Table 3-42 compares the range (and mean and median values) of FACx values for the 20 comparable JU03 and Urban 2000 predictions using these two protocols.

Table 3-42. Range [Mean/Median], Across Modes and MET Input Options, of FAC2 and FAC5/FAC10 Values (in parentheses) for 20 Sets of Urban HPAC Predictions of JU03 and Urban 2000 (MS-based JU03 predictions are not considered for these comparisons.)

Data Threshold	Field Trial / Condition	JU03	Urban 2000
Background + 1 ppt	**Day**	**0.44-0.56** [0.49/0.49] (0.70-0.86 / 0.82-0.94)	Not available
Background + 1 ppt	**Night**	**0.35-0.51** [0.42/0.40] (0.52-0.78 / 0.60-0.86)	**0.18-0.26** [0.23/0.23] (0.41-0.55 / 0.51-0.69)
15 ppt	**Day**	**0.21-0.35** [0.26/0.25] (0.55-0.74 / 0.71-0.89)	Not available
15 ppt	**Night**	**0.07-0.27** [0.17/0.16] (0.24-0.60 / 0.37-0.76)	**0.06-0.18** [0.13/0.13] (0.16-0.40 / 0.30-0.58)

Table 3-42 suggests a substantial improvement associated with the JU03 predictions relative to Urban 2000 (at night), perhaps due to improved MET inputs. When considering the lower threshold ("Background + 1 ppt"), and hence the edges of the predicted and/or observed plume, the JU03 nighttime FAC2 values (range, mean, and median) are about twice those of Urban 2000. For the 15 ppt threshold protocol, where the edges of the predicted and/or observed plumes might be excluded, there is still some

[14] We also examined a protocol that included all pairings, even comparisons in which both the observation and prediction were defined as the background. Finally, we also considered protocol that considered a 25 ppt threshold – that is, all comparisons in which both the prediction and observation were below 25 ppt were excluded. The conclusions discussed here were not changed by these additional analyses.

relative improvement associated with the JU03 FAC2 values (about 30%, e.g., [0.17-0.13]/0.13).

It was postulated during the analysis of the Urban 2000 field experiment and Urban HPAC predictions that the terrain associated with Salt Lake City represented a challenging and important feature that could influence the wind fields substantially. For example, a forecast that was used as MET input for the Urban 2000 study ("OMEGA") led to some of the best Urban 2000 predictions [Refs. 3-1 and 3-5] yet was recognized as missing the plume direction on the arcs.

In Chapter 2 of this paper, the use of the BGS MET input option was described as part of our initial investigations. This option was rejected for the final comparative analyses because it was determined to sometimes miss the wind directions and speeds significantly (see Figures 2-31 and 2-33). For the BGS-based Urban HPAC predictions, the range of FACx values (for the four non-MS modes) was:

- For the "background + 1 ppt" protocol
    - During the day for all surface samplers, FAC2 (5/10) range = **0.44-0.46** (0.74-0.75 / 0.86-0.86). FAC2 mean/median values were 0.45/0.45.
    - At night for all surface samplers, FAC2 (5/10) range = **0.18-0.25** (0.33-0.47 / 0.44-0.61). FAC2 mean/median values were 0.22/0.21.
- For the "15 ppt" protocol
    - During the day for all surface samplers, FAC2 (5/10) range = **0.21-0.24** (0.758-0.61 / 0.78-0.79). FAC2 mean/median values were 0.23/0.24.
    - At night for all surface samplers, FAC2 (5/10) range = **0.02-0.08** (0.16-0.30 / 0.30-0.49). FAC2 mean/median values were 0.05/0.05.

The nighttime FACx values associated with the BGS predictions are quite similar to those previously reported for Urban 2000 for the "background + 1 ppt" protocol. Therefore, it appears that a few missed wind directions, as was certainly the case for the nighttime BGS MET-based predictions, could account for the observed relative FAC2 results. Furthermore, the suggestion here is that improved MET input options that better represented the actual winds were available and used for the JU03 predictions, especially relative to the comparable 30-minute average concentration Urban 2000 predictions. Future efforts are planned to evaluate the latest version of Urban HPAC (including MSS) using the Urban 2000 field experiment. Such efforts may shed light on the relative differences discussed above.

REFERENCES

3-1. Warner, S., N. Platt, and J. F. Heagy, 2004: Comparisons of transport and dispersion model predictions of the URBAN 2000 field experiment, *J. Appl. Meteor.,* **43**, 829-846.

3-2. Warner, S., N. Platt, J. F. Heagy, J. E. Jordan, and G. Bieberbach, 2006: Comparisons of transport and dispersion model predictions of the mock urban setting test field experiment. *J. Appl. Meteor. and Climatology*, **45**, 1414-1428.

3-3. Warner, S., N. Platt, and J. F. Heagy, 2004: User-oriented two-dimensional measure of effectiveness for the evaluation of transport and dispersion Models, *J. Appl. Meteor.,* **43**, 58-73.

3-4. Heagy, J. F., N. Platt, S. Warner, and J. T. Urban, 2007: *JEM Urban IPT Final Report*, IDA Memorandum for DTRA/JSTO, April.

3-5. Warner, S. and N. Platt, 2003: *Analyses in Support of Initial Validation of Urban HPAC: Comparisons to Urban 2000 Observations,* IDA Document D-2870N. [Available from Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882.]

APPENDIX A

ACRONYMS

APPENDIX A
ACRONYMS

ACA	downwind ANL SODAR and Profiler observations used for a MET input option
A_{FN}	false negative region
A_{FP}	false positive region
AGL	above ground level
ANL	Argonne National Laboratory
A_{OB}	region associated with the observations
A_{OV}	region of overlap
A_{PR}	region associated with the predictions
ARA	Applied Research Associates
ARL	Air Resources Laboratory or Army Research Lab
ASTM	American Society for Testing and Materials
ASU	Arizona State University
BAS	Baseline MET input option
BG	Botanical Gardens
BGS	Botanical Gardens MET input option, Botanical Gardens mini-SODAR
BINEX	building interior model
BNMSE	Bounded Normalized Mean Square Error
BRB	GCAT MET input option
C	observed or predicted concentration or dosage
CALPUFF	a long-range transport air dispersion model
CBD	central business district
CBRN	chemical, biological, radiological, and nuclear
CC	Christian Church
CCAM	Conformal Cubic Atmospheric Model
CD	compact disc
CDT	Central Daylight Time
CFD	computational fluid dynamics
c-MOE	concentration-based MOE
CSIRO	Commonwealth Scientific and Industrial Research Organisation (Australia)
cm	centimeter
C_o	observed concentration or dosage
C_p	predicted concentration or dosage
CPU	Central Processing Unit

DHS	Department of Homeland Security
DLL	Dynamic Link Library
DM	UDM
DoD	Department of Defense
DPG	Dugway Proving Ground
DSTL	Defense Science and Technology Laboratory (United Kingdom)
DSTO	Defense Science and Technology Organization (Australia)
DTIC	Defense Technical Information Center
DTRA	Defense Threat Reduction Agency
DW	UDM and UWM
ETEX	European Tracer Experiment
FAC2	fraction of predictions within a factor of 2 of the observations
FAC5	fraction of predictions within a factor of 5 of the observations
FAC10	fraction of predictions within a factor of 10 of the observations
FACx	fraction of predictions within a factor of x of the observations
FB	Fractional Bias
FDDA	Four Dimensional Data Assimilation
FRD	Field Research Division
GCAT	Global Climatological Analysis Tool
GCM	GCAT MEDOC format
GMU	George Mason University
g/s	grams per second
HPAC	Hazard Prediction and Assessment Capability
IDA	Institute for Defense Analyses
IDL	Interactive Data Language
IOP	Intensive Operating Period
ITT	ITT Corporation
JEM	Joint Effects Model
JRO	Joint Requirements Office
JSTO	Joint Science and Technology Office
JU03	Joint Urban 2003 (field experiment)
kg	kilograms
km	kilometers
Lat	latitude
LiDAR (or LIDAR)	Light Detection and Ranging
Lo	low resolution
Log	logarithm

Lon	longitude
m	meters
max	maximum
MC-SCIPUFF	Mass-Consistent SCIPUFF
MEDOC	meteorological data format
MET	meteorological, meteorological input
MESO	a Monte Carlo Lagrangian dispersion code
MG	geometric mean bias
min	minutes
MINERVE	methode d'interpolation et de reconstitution tridimensionelle d'un champ de vent
MIRAN	Miniature Infrared Analyzer
MM5	Mesoscale Model (5^{th} generation)
MOE	measure of effectiveness
MS	MSS
MSS	Micro SWIFT SPRAY
MUST	Mock Urban Setting Test
n	number of data points
N	number of independent releases
NAD	Normalized Absolute Difference
NARAC	National Atmospheric Release Advisory Center
NCAR	National Center for Atmospheric Research
NGIC	National Ground Intelligence Center
NMSE	Normalized Mean Square Error
NOAA	National Oceanographic and Atmospheric Administration
Obs	observations
OMEGA	Operational Multiscale Environment Model with Grid Adaptivity
PA	Park Avenue
PIGS	Programmable Integrating Gas Samplers
PNA	upwind PNNL SODAR and Profiler observations used for a MET input option
PNNL	Pacific Northwest National Laboratory
PO7	Post Office (7^{th} processing technique) MET input option
ppt	parts per trillion
pptv	parts per trillion volume
Pred	predictions
PWIDS	Portable Weather Information and Display System
QUIC PLUME	a Los Alamos Lab-developed urban diagnostic model
QUIC URB	a Los Alamos Lab-developed Lagrangian dispersion model
R	linear Pearson correlation coefficient

RASS	Radio Acoustic Sounding System
Ref.	Reference
Refs.	References
R_{ln}	correlation coefficient based on first taking logarithms of values
RMSE	Root Mean Square Error
RTFDDA	real-time FDDA
RUSTIC	Realistic Urban Spread and Transport of Intrusive Contaminants
σ_C	standard deviation
s	seconds
SAIC	Science Applications International Corporation
SCIPUFF	Second-Order Closure Integrated Puff
SDF	Sensor Data Fusion
SF_6	sulfur hexafluoride
SODAR	Sonic Detection and Ranging or simply Acoustic Sounder
SONICS	sonic anemometer/thermometer
SPRAY	a Monte-Carlo Lagrangian dispersion code
STINET	Scientific and Technical Information Network
SWIFT	Stationary Wind Fit and Turbulence
T&D	Transport and Dispersion
U	University
UC	urban canopy
UDM	Urban Dispersion Model
UofU	University of Utah
UTC	Universal Time (Coordinated)
UTM	Universal Transverse Mercator
UWM	Urban Windfield Module
VG	geometric mean variance
VLAS	Variational LiDAR Assimilation System
VLSTRACK	Vapor, Liquid, and Solid Tracking
V&V	Verification and Validation
VV&A	Verification, Validation, and Accreditation
WA	Westin-A
WB	Westin-B
WindTracer	LiDAR used for measuring radial wind velocities
WM	UWM

APPENDIX B

SUPPLEMENTARY MOE FIGURES

APPENDIX B
SUPPLEMENTARY MOE FIGURES

This appendix provides supplementary MOE figures for Urban HPAC predictions of JU03. Twelve figures are presented. For each figure, the point estimates for the MOE values – i.e., the vector average for the 17 daytime or 12 nighttime releases – are shown. Each figure compares MOE values for 25 sets of Urban HPAC predictions – five modes (**UC**, **WM**, **DM**, **DW**, and **MS**) × five MET input options (BAS, BRB, PNA, ACA, and PO7). Point estimates are shown, as opposed to 99% confidence region as is the case in the main body of the report, so that all 25 MOE values can be easily seen in a single plot.

The first four figures examine 25 ppt threshold-based MOE values and the next four figures (Figures B-5 through B-8) show the 250 ppt threshold-based results. The final four figures present the concentration-based MOE results. Each set of four provides comparisons for the daytime/CBD, daytime/arcs, nighttime/CBD, and nighttime/arcs conditions, in that order. On several of the figures, we have highlighted model behavior that appears similar by simply drawing solid lines around clusters of values. These lines, and the enclosed clusters, do not have any special meaning in terms of statistical significance. Rather, this cluster is used to help the reader see some of the underlying trends, many of which have been discussed in the main body.

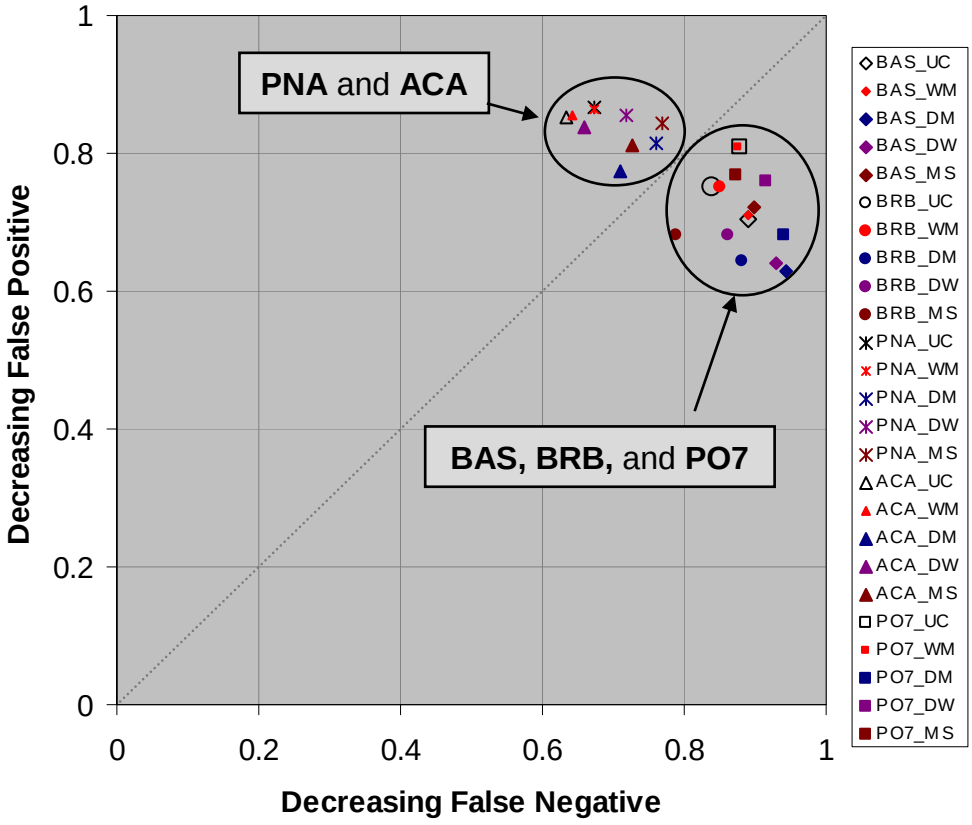


Figure B-1. Comparisons of 25 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Day Releases in the CBD – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

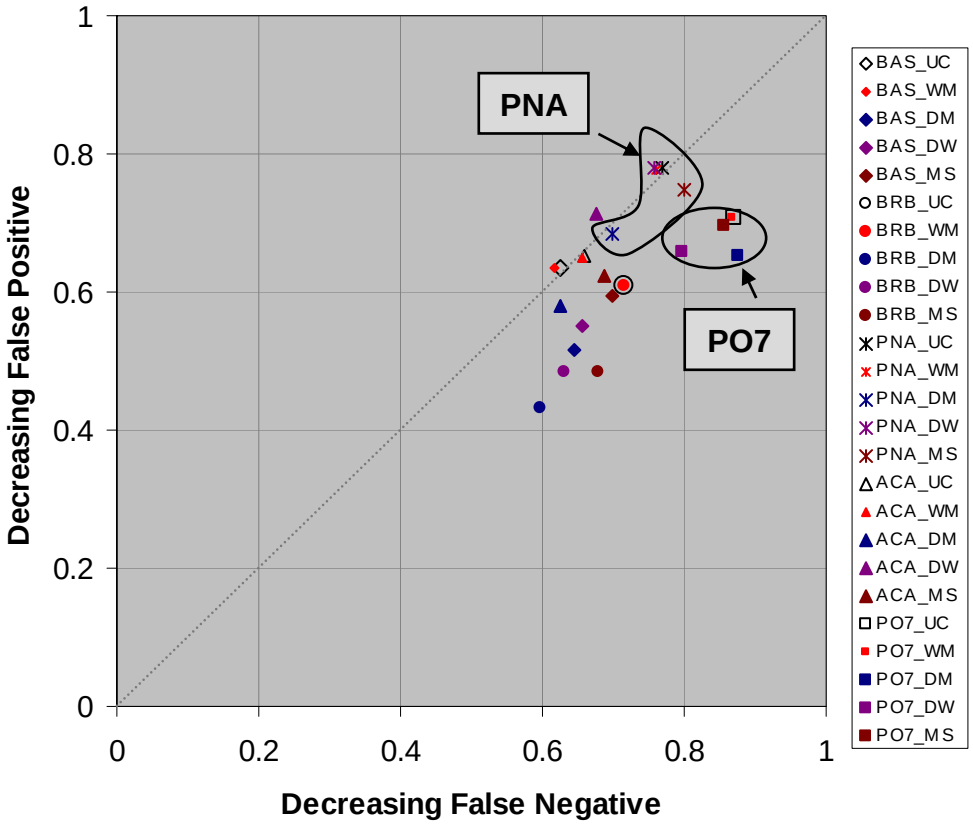


Figure B-2. Comparisons of 25 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Day Releases on the Arcs – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

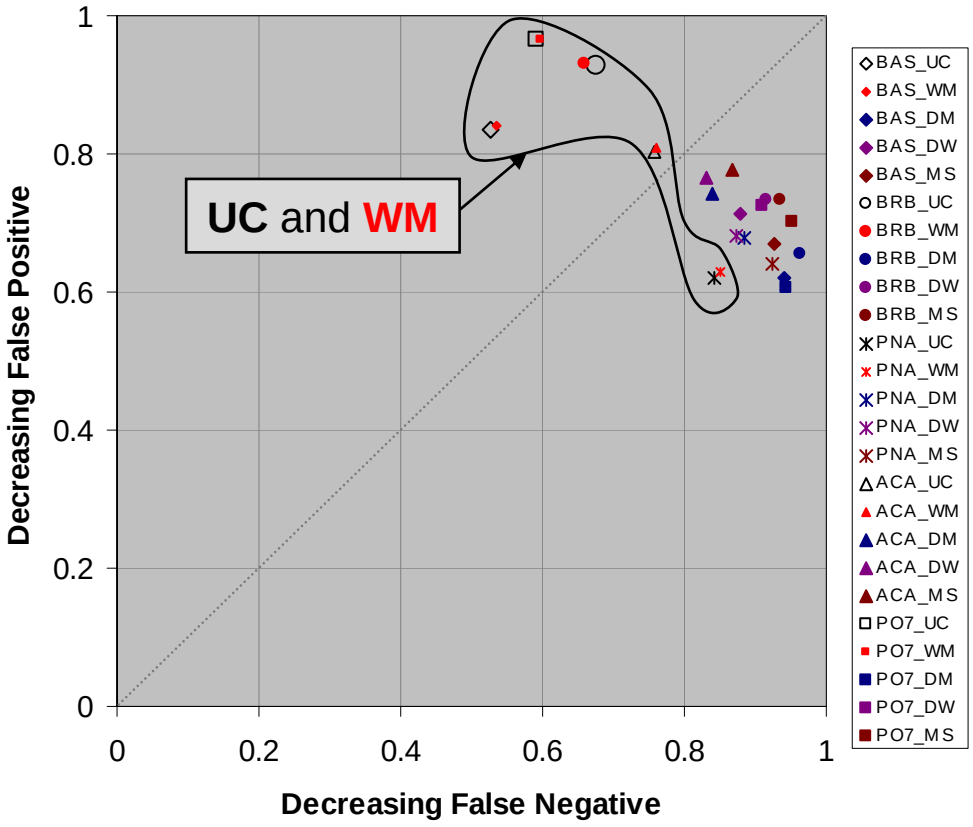


Figure B-3. Comparisons of 25 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Night Releases in the CBD – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

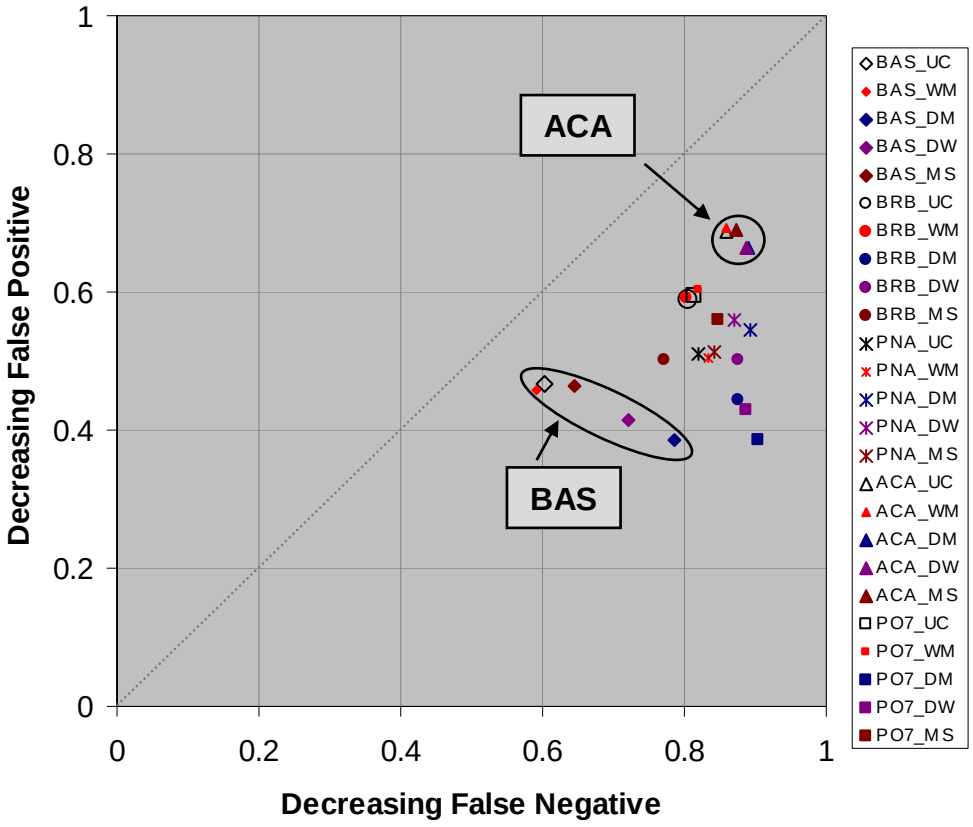


Figure B-4. Comparisons of 25 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Night Releases on the Arcs – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

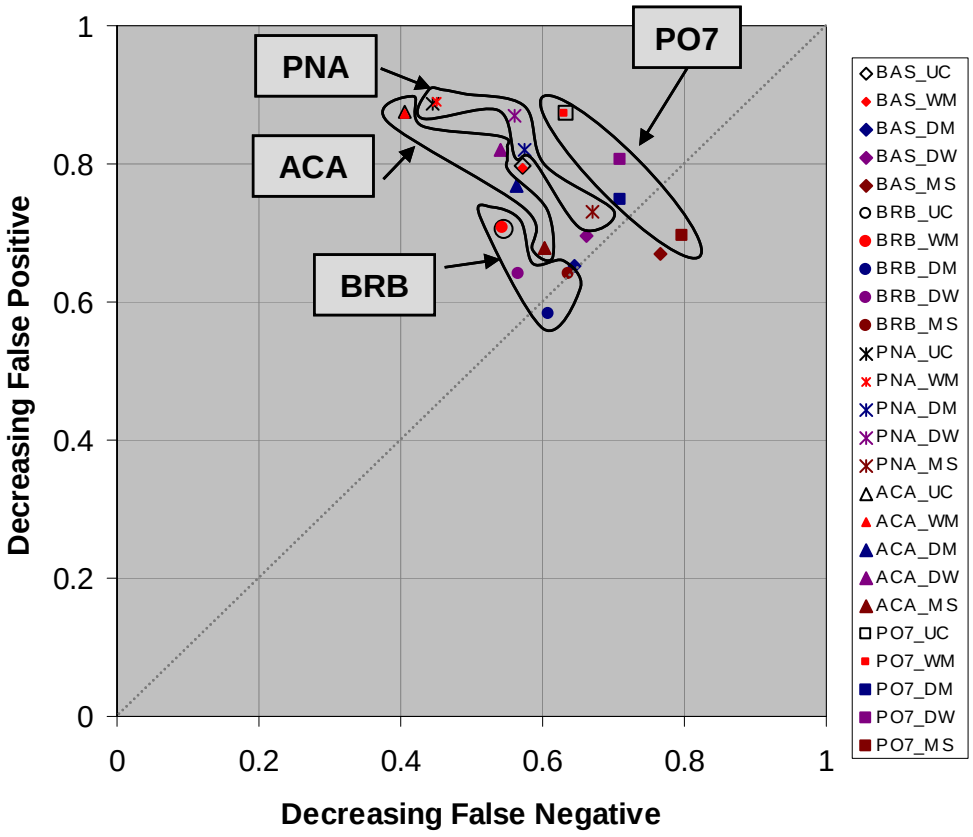


Figure B-5. Comparisons of 250 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Day Releases in the CBD – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

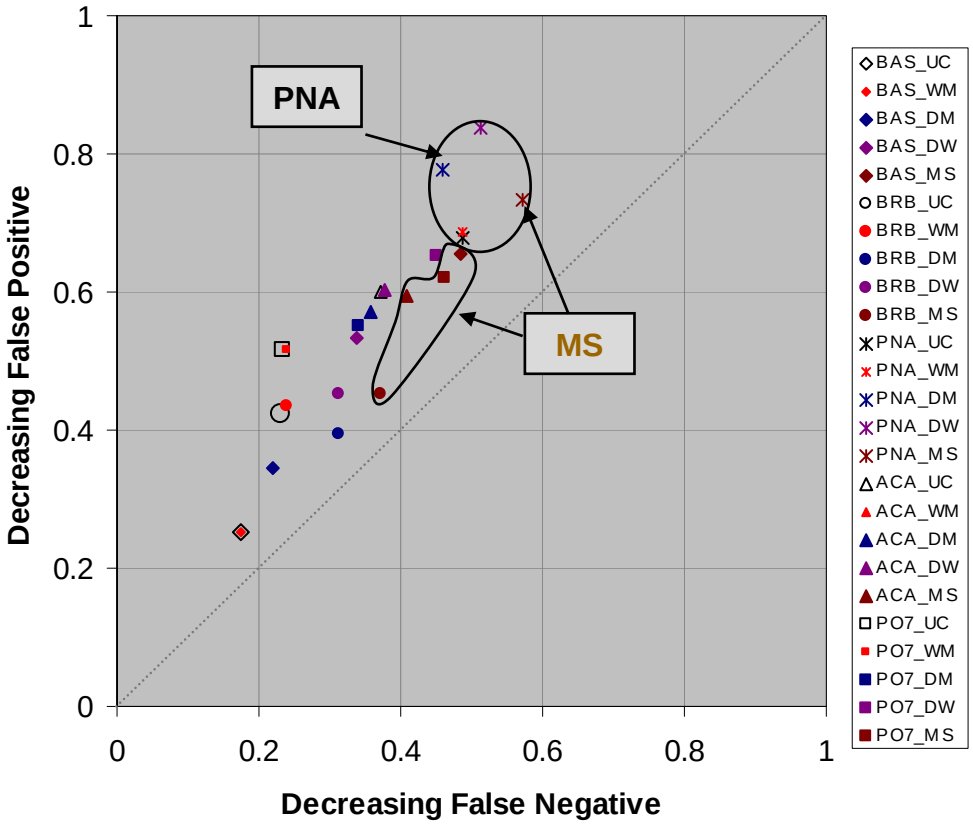


Figure B-6. Comparisons of 250 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Day Releases on the Arcs – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

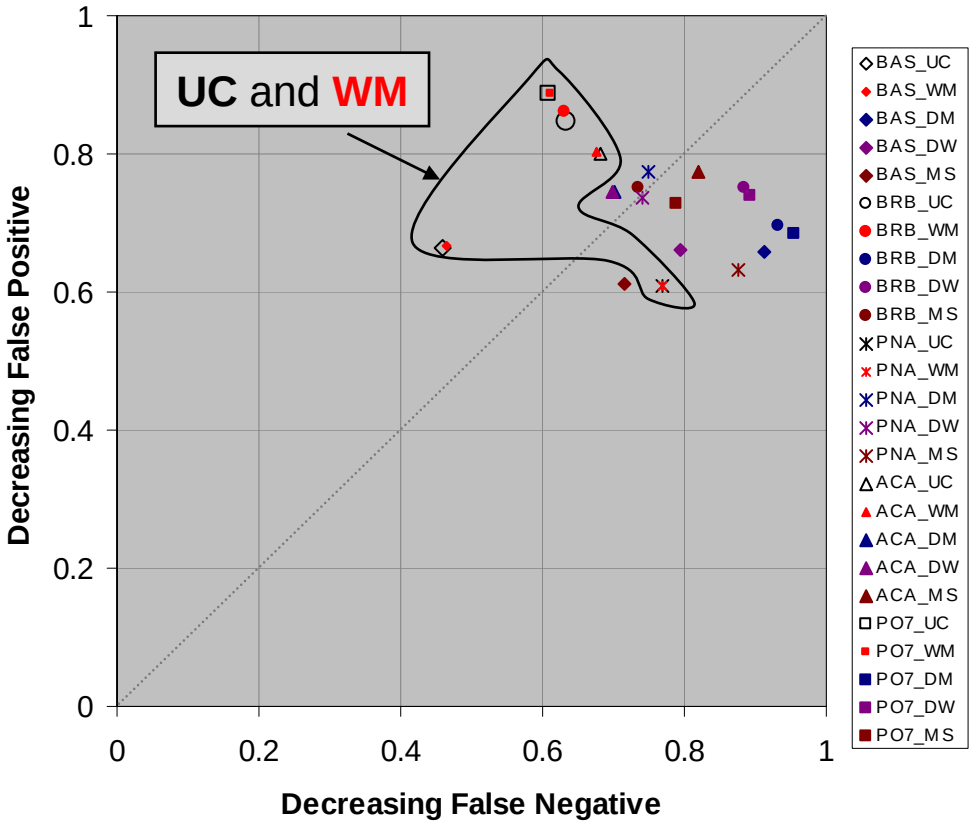


Figure B-7. Comparisons of 250 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Night Releases in the CBD – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

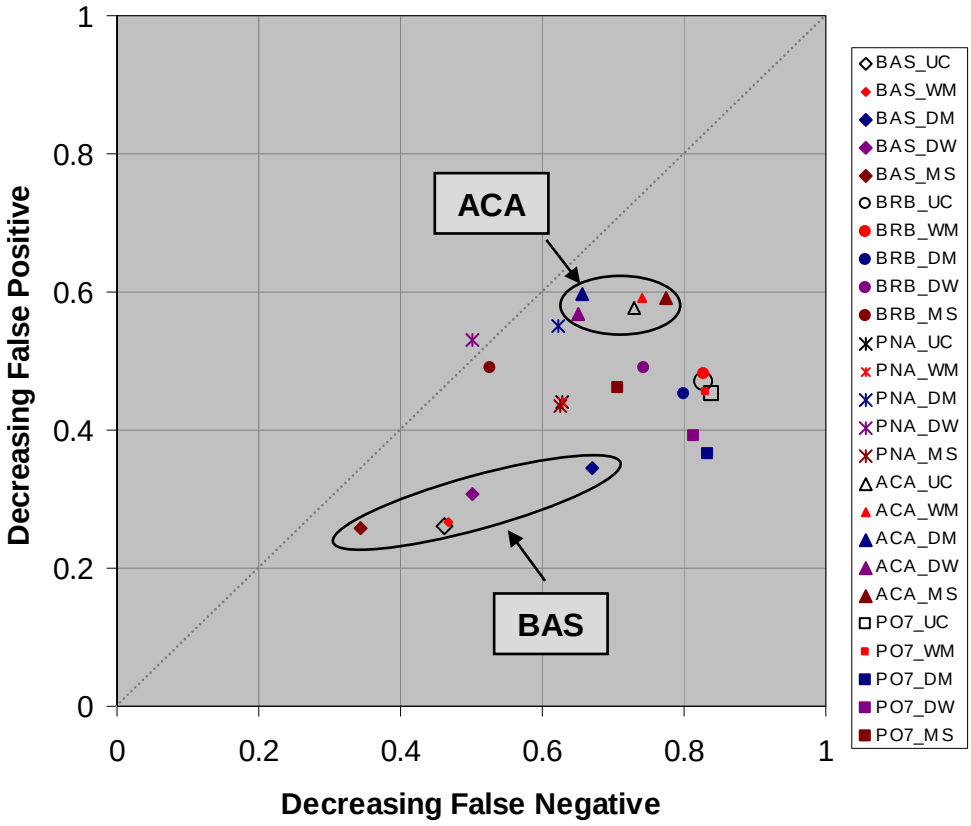


Figure B-8. Comparisons of 250 ppt Threshold-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Night Releases on the Arcs – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

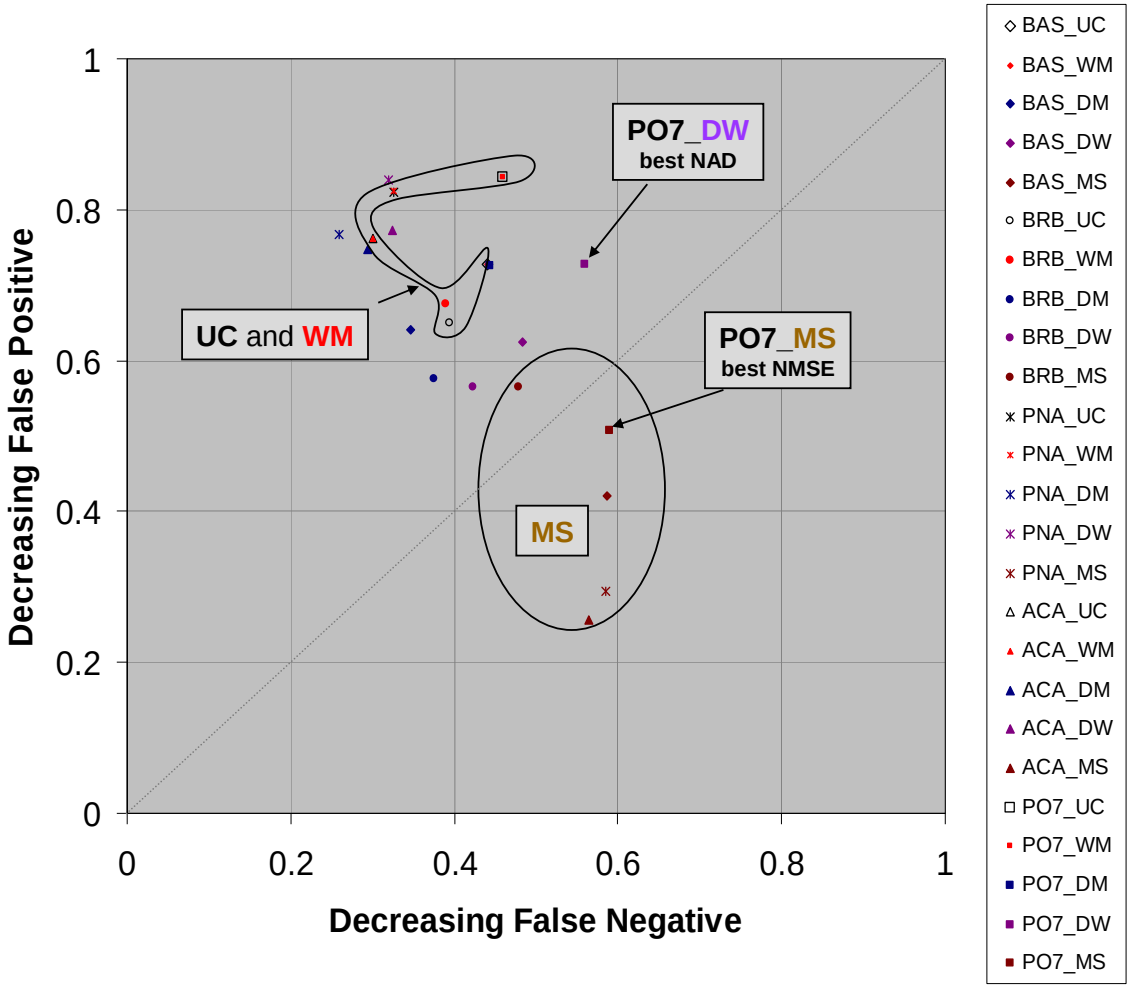


Figure B-9. Comparisons of Concentration-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Day Releases in the CBD – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

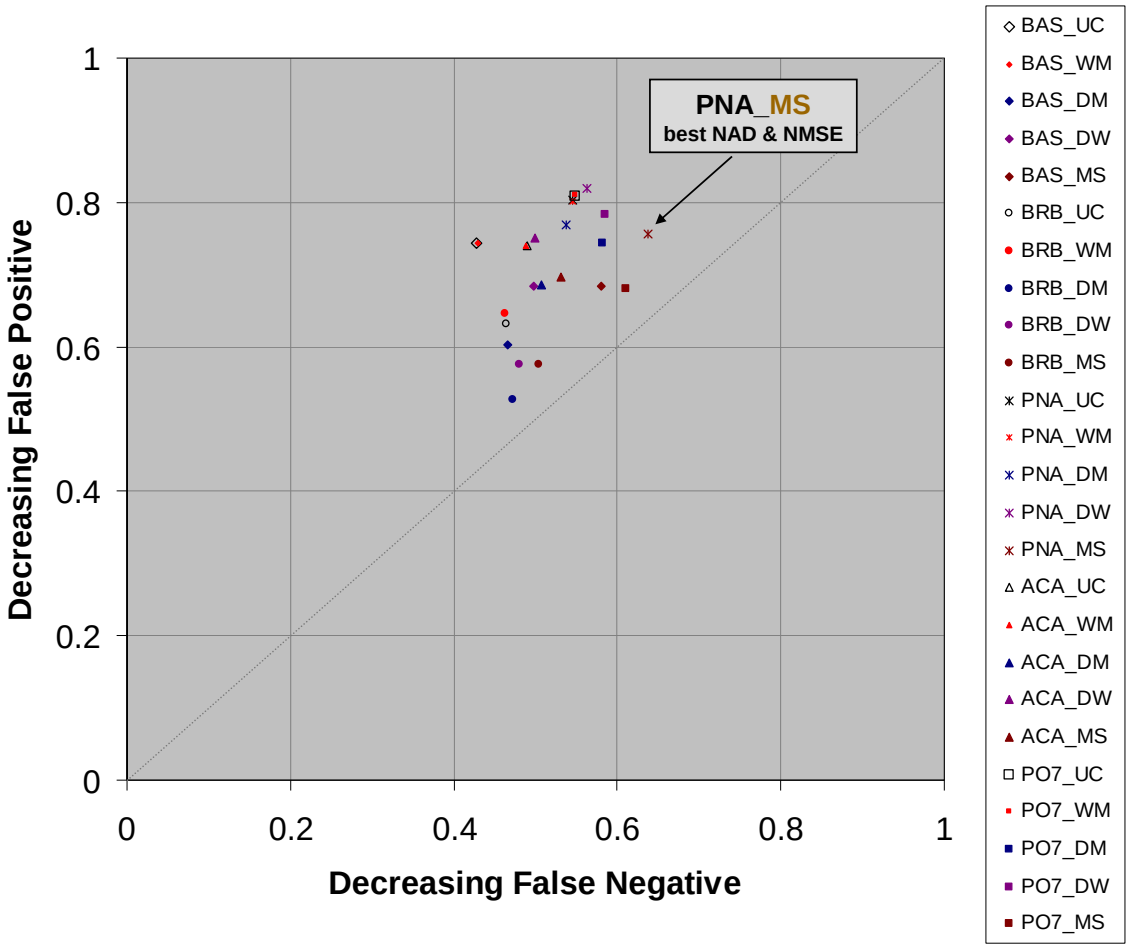


Figure B-10. Comparisons of Concentration-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Day Releases on the Arcs – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

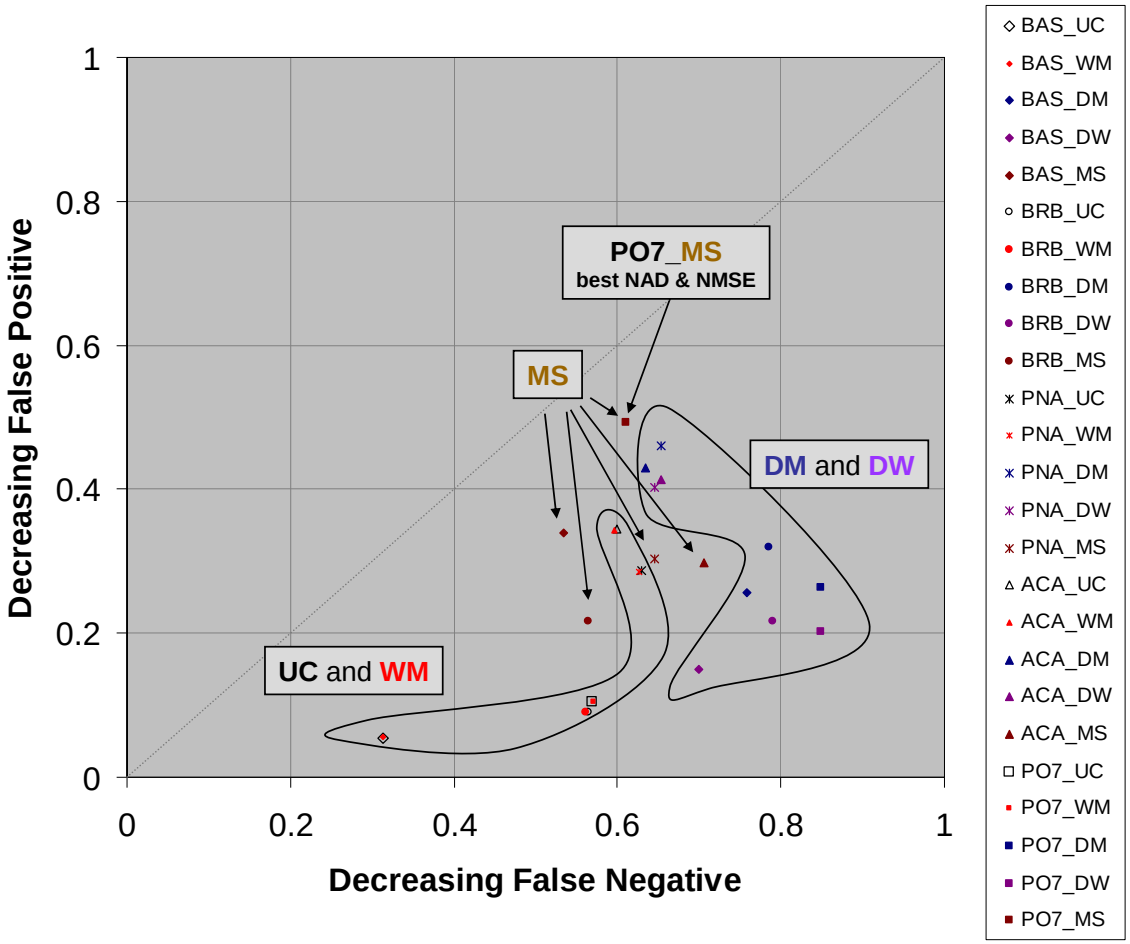


Figure B-11. Comparisons Concentration-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Night Releases in the CBD – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

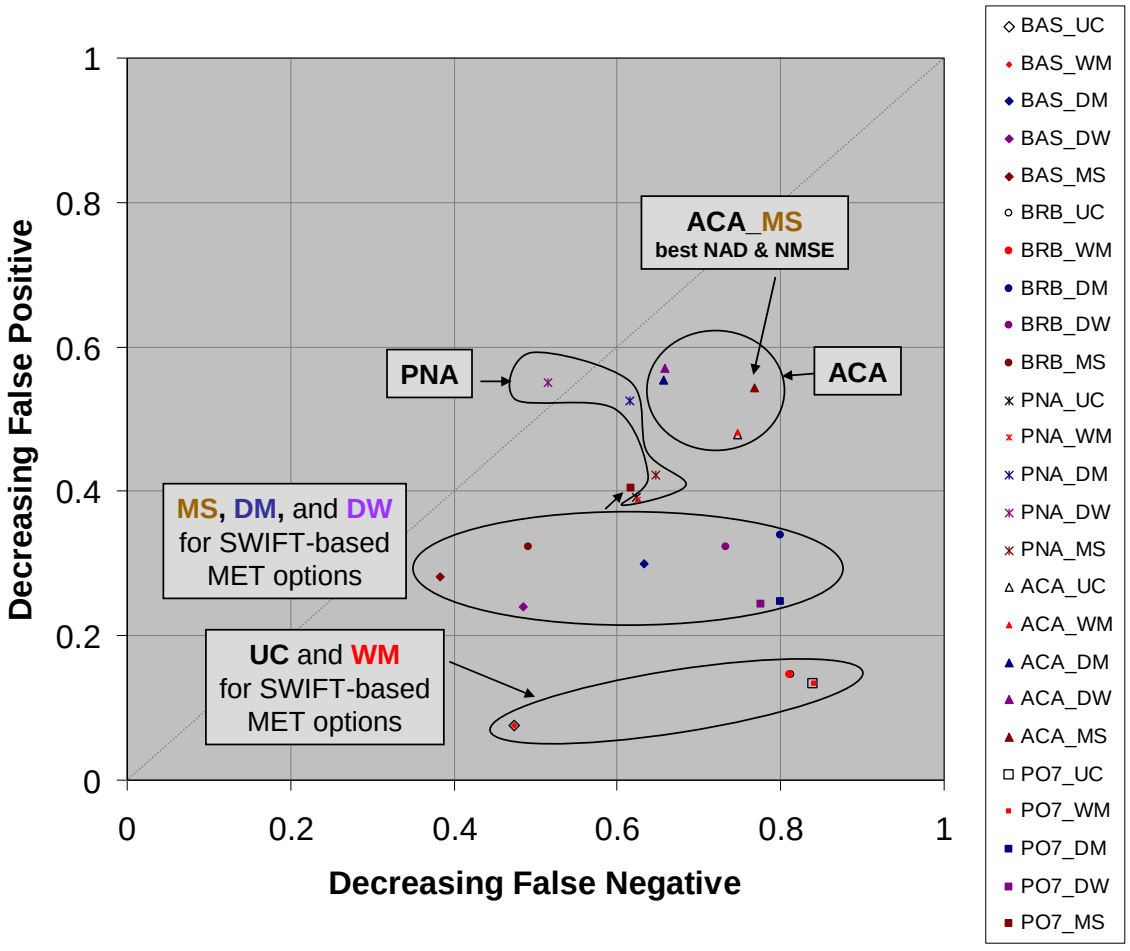


Figure B-12. Comparisons of Concentration-Based MOE Values for 25 Sets of Urban HPAC Predictions of JU03 Night Releases on the Arcs – Five Modes (UC, WM, DM, DW, and MS) × Five MET Input Options (BAS, BRB, PNA, ACA, and PO7).

APPENDIX C

TASK ORDER EXTRACT

APPENDIX C
TASK ORDER EXTRACT
(WITH MOST PERTINENT SECTION IN RED FONT)

DC-1-2607

TITLE: Support for DTRA in the Validation Analysis of Hazardous Material Transport and Dispersion Prediction Models

This task order is for work being performed by the Institute for Defense Analyses (IDA) under Contract Number DASW01-04-C-0003/W74V8H-05-C-0042 for the Defense Threat Reduction Agency (DTRA).

1. BACKGROUND:

The DTRA/Joint Science and Technology Office (JSTO) Verification and Validation (V&V) Program represents ongoing activities performed in parallel with development of all predictive codes in support of hazardous material transport and dispersion prediction. One element of V&V is to perform code-on-code comparisons. In this strategy, each code receives the same input. In this manner, differences in the output predictions can lead to the identification of software bugs, or help to assess technical strengths and weaknesses of component algorithms within each code. In addition, a certain amount of credibility for both models is achieved when their predictions agree. When the inputs are simple, such as for fixed winds and simple terrain, the predictions tend to be dominated by the dispersion algorithms. Comparisons at this level of complexity are important to establish fundamental dispersion algorithm veracity, and to help discover software bugs. As more complex terrain, an urban environment, and weather are included as input, the number of physical processes responsible for transport and dispersion increases and the predictions become the result of many interdependent algorithm calculations.

It is very difficult to separate meteorological uncertainty from the transport and dispersion model accuracy when comparing predictions to field-trial validation quality or real-world data. The validation challenge is to assess whether a model performs well over different field trials, and ultimately reflects real-world phenomena. Some codes perform better under certain conditions and specific scenarios. Hazard prediction models are generally developed for a range of user communities and applications. Each user community has a different set of requirements. Thus, the corresponding hazard models tend to be optimized for specific applications. The process of validating a model should be couched in terms of end-user requirements where feasible.

Various metrics are used to express model performance relative to observed data. Most metrics tend to use manifestations of a ratio (geometric or arithmetic) between the predicted and measured quantities. The compared quantities are usually peak, plume-centerline, and off-axis concentration or dosage, as well as crosswind and along-wind spread and area coverage. Other metrics might include the second moment of the dosage and concentration values at a sampler location. All these metrics are reasonable validation performance measures, but none of them explicitly expresses an application-oriented performance measure. A "yardstick" that measures application-oriented model performance is needed. The scale on this yardstick would clearly and directly relate to the specific user's concerns and needs. The pursuit (and application) of "validation" performance measures (Measure of Effectiveness or MOE) is a continuing initiative.

Several aspects of hazard prediction modeling are the subject of current improvement programs. First, complexities associated with the urban environment are being addressed via an urban transport and dispersion program. Codes varying from empirical (wind tunnel-based) to computational fluid dynamics-based are being considered to address the complex flows associated with an urban environment. As they become mature (and validated), tools to address the infiltration, exfiltration, and flow within buildings and other complex structures are also being considered for inclusion within hazard prediction models. Algorithms to estimate source term parameters (e.g., location, time, and amount) from sparse observations are also being developed. Such "sensor data fusion" tools are expected to improve hazard predictions in scenarios where the release is covert or accidental. Field experiments have been conducted, and are being designed, to aid in the evaluation of urban (including within building) and sensor data fusion models. These evaluations are crucial to the overall management of these programs.

2. OBJECTIVE:

IDA will conduct independent analyses and special studies associated with verification, validation, and evaluation of the suite of models associated with the Hazard Assessment and Prediction Capability. IDA will support development of user-oriented performance MOEs using field trial data sets and will coordinate scenario definition and arbitration for code-on-code V&V activities.

The objectives of these analysis and coordination are (1) to ensure that a consistent analysis approach is used when comparing model predictions, and assist DTRA in the implementation of code-on-code analysis, comparisons, and interpretation; and (2) to define measures of effectiveness in terms of user-specific objectives and applications. The scope of this effort may be expanded to other programs as directed by DTRA.

3. STATEMENT OF WORK:

As required by DTRA, IDA will perform the following tasks:

a. Comparisons of DTRA-Identified Urban and Building Interior T&D Models. IDA will continue comparisons of urbanized versions of HPAC to field trial experiments. These comparisons will include the creation of multiple sets of Urban HPAC predictions (to be used for sensitivity and diagnostic studies) and include at least one set of

operationally-oriented predictions. Qualitative (graphical) and quantitative comparisons, including the use of standard statistics and the user-oriented MOE, will be conducted. As a minimum, the Joint Urban 2003 field trial will be used for Urban HPAC (to include Micro SWIFT/Micro SPRAY - MSS) comparisons. IDA will also examine the impact of the incorporation of various meteorological inputs on the quality of model predictions. A re-examination of the Urban 2000 field trial may also be conducted using future versions of Urban HPAC (included the latest urban windfield model and MSS).

Given an integrated HPAC building interior model (e.g., BINEX), comparisons of predictions of the transport and dispersion of pollutants within buildings will also be studied. Other urban field trials or wind tunnel experiments to be identified by IDA (with DTRA concurrence) or DTRA will also be considered.

b. Joint Validation Studies. As part of the Joint Urban 2003 studies, comparisons between Urban HPAC, MESO/RUSTIC (ITT), and QUIC URB / QUIC PLUME (Los Alamos) will also be conducted. As Joint Effects Model (JEM) versions become available, it is expected that similar comparative studies will be conducted.

IDA may continue to conduct quantitative controlled comparisons between DoD's HPAC and DHS's NARAC (if joint collaborative efforts are identified). Other collaborative arrangements identified by IDA (with DTRA concurrence) or DTRA to examine comparisons with other T&D models may also be considered.

c. Sensor Data Fusion (SDF) Related Studies. As directed by DTRA, IDA will provide technical and analytical support associated with the initial incorporation of SDF algorithms into hazard prediction tools and products. This new area will likely require the development of new analytical techniques and procedures/protocols in order to support credible quantitative assessments of this emerging technology area.

d. Missile Intercept Modules. IDA will provide technical support to expected initiatives to improve the modeling and integration of modules that characterize the source terms associated with a release based on a missile intercept. Efforts will include support in identifying requirements and evaluating solutions.

As a part of the all of the above subtasks, IDA will communicate, via conference papers and/or posters, working group discussions, and IDA papers, the more important applications of the MOE and any progress toward the creation of "demonstration" validations. In addition, IDA should create descriptions of its efforts, where appropriate (and approved by DTRA), that are suitable for publication in peer-reviewed journals. IDA will actively participate in working groups (e.g., Sensor Data Fusion) and Technical Panel 9 as directed by DTRA.

4. CORE STATEMENT:

This research is consistent with IDA's mission in that it will support specific analytical requirements of the sponsor and will assist the sponsor with planning efforts. Accomplishment of this task order requires an organization with experience in operationally oriented issues from a joint and combined perspective, which IDA, a Federally Funded Research and Development Center, is able to provide. It draws upon IDA's core competencies in Systems Evaluations and Operational Test and Evaluation.

Performance of this task order will benefit from and contribute to the long-term continuity of IDA's research program.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YY)	2. REPORT TYPE	3. DATES COVERED (FROM – TO)
March 2007	Final	October 2005–March 2007

4. TITLE AND SUBTITLE

Comparisons of Transport and Dispersion Model Predictions of the Joint Urban 2003 Field Experiment

5A. CONTRACT NO.

DASW01 04 0003/W74V8H-05-C-0042

5B. GRANT NO.

5C. PROGRAM ELEMENT NO(S).

6. AUTHOR(S)

Warner, S. ,Platt, N., Urban, J.T., and Heagy, J.F.

5D. PROJECT NO.

5E. TASK NO.

DC-1-2607

5F. WORK UNIT NO.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

8. PERFORMING ORGANIZATION REPORT NO.

IDA Paper P-4195

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Threat Reduction Agency
Joint Science & Technology Office for Chemical & Biological Defense
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6201

10. SPONSOR'S / MONITOR'S ACRONYM(S)

DTRA / Joint Science and Technology Office (JSTO)

11. SPONSOR'S / MONITOR'S REPORT NO(S).

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The potential effects of the atmospheric release of hazardous materials continue to be of concern to the nation. This concern is especially acute in urban areas where denser populations would be exposed to hazardous materials. In addition, military activities increasingly are conducted in urban, highly populated areas. Therefore, the release of hazardous materials in an urban environment is troubling to both civilian and military authorities. Improved characterization and understanding of urban transport and dispersion will allow for more robust modeling. Under the joint sponsorship of the DoD (DTRA) and DHS, a series of tracer gas releases were carried out in Oklahoma City in the summer of 2003, and this field experiment is referred to as "Joint Urban 2003." For this study, several sets of predictions based on different configurations of Urban HPAC and varying meteorological inputs were created, and the resulting concentration predictions were compared with the Joint Urban 2003 observations. In addition, comparisons between sets of Urban HPAC predictions are examined, and thus, relative performance among Urban HPAC configurations is assessed. These evaluations are also briefly compared with previous related studies.

15. SUBJECT TERMS

model validation; hazardous material transport and dispersion; HPAC; urban dispersion; measure of effectiveness; Joint Urban 2003

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NO. OF PAGES	19A. NAME OF RESPONSIBLE PERSON
A. REPORT	B. ABSTRACT	C. THIS PAGE	Unlimited	244	Dr. John Hannan
Unclassified	Unclassified	Unclassified			19B. TELEPHONE NUMBER (INCLUDE AREA CODE) 703-767-3286